USING AUTO-ORDERING

TO IMPROVE OBJECT TRANSFER BETWEEN MOBILE DEVICES

A Thesis Submitted to the College of

Graduate Studies and Postdoctoral Studies

In Partial Fulfillment of the Requirements

For the Degree of Master of Science

In the Department of Computer Science

University of Saskatchewan

Saskatoon, CANADA

By

Chengzhao Li

PERMISSION TO USE

Head of the Department of Computer Science
176 Thorvaldson Building
110 Science Place
University of Saskatchewan
Saskatoon, Saskatchewan
Canada
S7N 5C9

ABSTRACT

People frequently form small groups in many social and professional situations: from conference attendees meeting at a coffee break, to siblings gathering at a family barbecue. These ad-hoc gatherings typically form into predictable geometries based on circles or circular arcs (called F-Formations). Because our lives are increasingly stored and represented by data on handheld devices, the desire to be able to share digital objects while in these groupings has increased. Using the relative position in these groups to facilitate file sharing could facilitate intuitive interfaces such as passing or flicking. However, there is no reliable, lightweight, ad-hoc technology for detecting and representing relative locations around a circle. In this thesis, we present three systems that can auto-order locations about a circle based on sensors standard on commodity smartphones. We tested two of these systems using an object-passing task in a laboratory environment against unordered and proximity-based systems, and show that our techniques are faster, more accurate, and preferred by users.

ACKNOWLEDGMENTS

I would like to express my gratitude to my supervisors Carl Gutwin and Kevin Stanley for their guidance and input throughout my graduate career at the University of Saskatchewan. The past two years was a memorable journey of my life. I am sincerely thankful for everything they have done for me. I also want to thank the faculty and staff in the Department of Computer Science for being great at what they do, as well as all of the great members of the HCI Lab for their friendship and support.

CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

AR	Augmented Reality
F-formation	Facing Formation
HCI	Human Computer Interaction
NFC	Near-Field Communication
RSSI	Received Signal Strength Indicator
OpenCV	Open Source Computer Vision

CHAPTER 1

INTRODUCTION[1]

People commonly need to transfer objects and files from one mobile device to another. For example, conference attendees talking at a coffee break might decide to share business cards or research papers; a family gathered in a living room might share photos of vacations or grandchildren; and colleagues sitting around a conference table might need to share files or data related to a project. Typically, individuals will arrange themselves into circular or semi-circular physically proximate locations know as Facing-formations [37, 38, 46] (F-formations). F-formations are the spatial patterns used by people involved in face-to-face interactions.

Although files and photos are now ubiquitously available to people on their mobile devices, transferring an object to another device can be cumbersome. Current file transfer techniques (such as e-mailing the object or creating a link to a location in the cloud) require time and information, such as e-mail address and online link address. Instead, transfer could be accomplished using techniques that allow people to move objects by bumping the devices [30], making parallel gestures [31, 32], or providing a list view of all connected devices [25]. However, these techniques have limitations – for example, bumping requires close physical proximity, parallel gesture can take time to decide on and execute, and lists require that people know the mapping between the list entry and the device in the real world.

Another class of technique uses onscreen targets for transferring objects from one device to another [24, 50]. Target-based object transfer is lightweight and natural compared to traditional file-sending mechanism, because people can simply direct the object of interest toward the target. There are two main aspects to a target-based object transfer technique: the local gesture used to select and direct the object, and the underlying infrastructure that identifies the target of the gesture. In this thesis, we are interested in the second of these issues – identifying target locations.

[1] Portion of the content in this thesis originally appeared in the following publication:
Li, C., Gutwin, C., Stanley, K., Nacenta, M.A. (2016). All Across the Circle: Using Auto-Ordering to Improve Object Transfer between Mobile Devices. In *Graphics Interface* (GI'2016), Victoria, BC, Canada. 49-56.

1.1 PROBLEM

The problem to be addressed in this thesis is that there is a lack of simple, intuitive and reliable localization techniques for determining device ordering.

File transfer between physically proximate individuals is easy in the physical world. People just need to get close enough to the targets and hand the physical files to them. In the electronic world, sharing a file can be far more cumbersome. Using e-mail or Dropbox to transfer files requires that people know additional information (email address or Dropbox link) before performing the file transfer task. Researchers have investigated using onscreen targets for transferring objects. In previous work, the target-identification problem has usually been solved employing approaches that require considerable infrastructure, such as magnetic sensors, infra-red cameras, multiple fiducial markers, or depth cameras [39]. There are many real-world situations where the difficulty of solving the target-location problem limits object transfer techniques. Although a few infrastructure-free approaches have been developed (e.g., the Virtual Compass [6]), these have high spatial error rates, which make transfers in small circular groups infeasible.

Most previous researchers have attempted to represent the precise target spatial location of user's devices to a centimeter level accuracy, a very difficult open problem which required extra infrastructure to address. This requirement makes it more difficult for their solutions to be deployed in real-life scenarios. For example, solutions using infra-red camera (e.g., the Vicon system [65]) to track targets require several cameras to be mounted in the room. This is obviously impractical for the ad hoc scenarios outlined above.

Meanwhile, the few infrastructure-free techniques require specific highly structured environments. For example, some techniques use trilateration with Wi-Fi signals to determine location. However, this method requires many adequate routers and has limited accuracy and precision. Approaches leveraging the inaudible stereo sound requires close relative distance. Chapter 2 will discuss the weaknesses of these current solutions.

1.2 MOTIVATION

The main motivation for developing a localization technique for device ordering is to improve the efficiency of data-sharing. Frequent data exchange occurs almost every day of our lives. Having an intuitive positioning technique can effectively reduce the difficulty of sharing data in ad-hoc circumstances, which would be beneficial to many people. For example, for a group of people sitting around for a lab meeting, it would be intuitive to send experiment results to group members by pointing the tablet, which contained the research report, at the people who needed it. A simple and quick localization technique that could provide sufficient accuracy and precision to allow for reliable real world aiming could address this problem.

In addition, lightweight localization techniques not only benefit the object transfer task but could also be useful in other fields. For instance, multiplayer modes are commonly available in many games. When players' relative locations are available as an input, simple matchmaking and novel game mechanism become possible.

1.3 SOLUTION

The solution explored in this thesis is to provide simple, intuitive and reliable localization techniques by using sensors standard on commodity mobile devices. We developed three systems that can auto-order locations about a circle based on different sensors commonly available on today's smartphones. Our solution draws on the idea of F-formation which describes the physical arrangement that people adopt when they come together to form a group. The constraint of F-formation geometry helps us narrow the problem from detecting precise locations in a plane to calculating relative locations about a circle.

The first of our auto-ordering techniques uses the smartphone's camera to find a fiducial marker (on paper or displayed on one of the phones themselves) placed in the middle of the group, and then uses the relative orientation of the marker observed by each phone to infer relative location.

The second technique uses the smartphone's compass: users orient their phones toward a location at the center of the circle, and the relative angle each device reports with respect to magnetic North can be used to resolve their positions.

In the third technique, each phone's camera measures the brightness gradient of the background; this gradient gives an absolute direction that can be shared with the other devices to determine a relative angular ordering.

Beyond these three systems, we also investigated the possibility of the combination of these solutions. For example, a Bayesian sensor fusion version, which combines compass and light gradient techniques, has been developed.

1.4 STEPS IN THE SOLUTION

In order to achieve the goal of the proposed solution, six steps were completed during the research process.

1.4.1 Step One: Understand Previous Research

It is important to review previous studies that focused on determine device location. There are a number of solutions available, but most of them require external infrastructure. This shortcoming implies that their solutions cannot be easily implemented by current commodity hardware. Our goal was to provide a solution by using sensors that are available on most of current mobile devices. Although some of the existing techniques do not require external equipment, they do have other shortcomings, such as low accuracy and precision, or distance limitations. We investigated the strengths and weakness of current studies and addressed them in our solutions.

1.4.2 Step Two: Constraint of F-formation

Our solutions were built on the idea of F-formations. A social group can be interpreted as a unit composed of several people who stand in a pattern with specific spatial relationship among them. Adam Kendon [37, 38] examined these patterns and described circular clusters called Facing Formations (F-formations) which described the physical arrangements that people adopt when they engage in focused conversational encounters. F-formations have been applied in previous technologies (e.g., [43, 45, 69]). The purposes of using F-formation in this thesis is to use the physical arrangement patterns of the F-formation system to constrain the localization problem. We collected information that could be extracted from given patterns and applied them to our solutions.

After systematically analyzing the characteristics of F-formations, we must define the factors that can be used to determine the ordering of devices. As we have simplified the problem from finding precise locations to calculating relative orientations, we needed to determine attribute or parameter that represents a unique location in a specific coordinate system. The constrained coordinate system is defined by the pattern described in F-formations. Our ordering techniques, described in Chapter Three, identify different factors for different technique solutions and test them in the last step.

1.4.3 Step Three: Sensor-based Auto-ordering Techniques

The third step was to develop solutions based on the design framed around F-formations. Under the constraint of F-formations, we developed three positioning systems. The first system takes the advantage of augmented reality technology. A fiducial marker, commonly used in simple Augmented Reality (AR) games, is used to provide a center point of a polar coordinate system. Devices' relative orientations are calculated by their positions in this coordinate system. Our second technique uses the device's orientation sensors. The constraint of F-formations allows us to create a polar coordinate system once all devices are pointing at the center of the grouping. The third system employs image processing techniques. The light gradient on a surface of a table or floor can provide a reference vector for forming a coordinate system. Device positions in this system can be inferred by calculating the light direction acquired by the equipped camera on the device, then comparing the relative orientations.

We also made some extended work based on our auto-ordering techniques. We developed a file transfer application on the Android platform. By using this application, a file on a mobile device can be easily selected and sent to other devices in the same group. We also developed a lock mechanism which relaxed constraints on users to remain in a particular position. The combination of our auto-ordering techniques was investigated as well in this step.

1.4.4 Step Four: Evaluation of Auto-ordering Techniques

We applied two of our solutions in an object transfer task to evaluate the performance of our solutions. We also employed two standard instruments to compare the performance and user preference of these techniques. Participants were asked to transfer several objects to others in the group. Both the objective and the subjective data collected in this experiment indicate that our solutions are superior to traditional techniques from both a performance and preference perspective.

1.5 EVALUATION

The techniques proposed in this thesis were subjected to both technical and usability assessments.

1.5.1 Technical Assessment

For the technical assessment, we evaluated the sensitivity, precision and span of the localization techniques, which showed that our solutions could provide sufficient accuracy and precision to reliably localize people around a circle.

1.5.2 Usability Assessment

Usability was evaluated by conducting a controlled study to compare our techniques to a proximity-based technique and an un-ordered portal technique. Participants were asked to perform several rounds of object transfer task using different location determining techniques. We measured two performance metrics (time and error rate) to determine if our solutions had better performance compared to two existing techniques, and asked subjective questions of the participants (NASA Task Load Index [26] and general preference questionnaire) to determine their subjective experience.

1.6 FINDINGS AND CONTRIBUTIONS

Our evaluation found:

- Object transfer with the auto-ordering techniques was faster than the other techniques, and less error prone than the portal technique.
- While participants were able to reduce transfer time using portals as they learned mappings, they never achieved better performance than the auto-ordering techniques.
- Participants overwhelmingly preferred the auto-ordering techniques to the portal and proximity techniques.
- The smartphone sensors underlying the techniques are accurate enough for groups of up to twenty people – many more than will typically be encountered in ad-hoc groups.

Our techniques provide a simple, intuitive and reliable solution for a common transfer situation — a small group gathered in an approximate circle — with sufficient accuracy and precision to reliably localize people around a circle.

This thesis makes three primary contributions: we have developed new techniques for effective object sharing, but in doing so have also added to the literature by providing comparative analysis and theoretical validation.

In this thesis, we developed three new localization techniques for determining device ordering. Our auto-ordering techniques are based on sensors and computational resources readily available on almost all smartphones.

Through the comparative evaluation, we provide empirical evidence concerning the tradeoffs between the different techniques for mobile application designers. For example, the un-ordered portal-based technique is suitable for repetitive object transfer tasks. Our techniques would be preferred in most of the cases discussed in the research of Marquardt et al. [46].

Our design draws heavily on the idea of F-Formations, which shows the spatial arrangements that people typically adopt in ad-hoc groups. Our research provides additional support to the validity and utility of F-Formations as a construct for designing co-located collaborative systems.

Beyond our immediate contributions, this thesis could have significant impact on other areas. For example, the techniques described in this thesis could work as an interface widget in other file sharing studies. Our technology demonstrates how simple spatial sensors can be used in clever ways to facilitate collaborative actions such as player interaction in co-located games.

1.7 THESIS OUTLINE

The thesis is organized as follows:

Chapter Two reviews previous works in several areas that form the foundation for the research presented in this thesis. First, current solutions for determining mobile devices' locations are discussed. We divide them into five major kinds: infrared/ultrasound, Wi-Fi/Bluetooth beacons, motion capture, vision technology and sensors on mobile device. After that, we present the theoretical grounding of our solution – F-formation and review the techniques developed by others based on this concept. Third, we summarize object transfer techniques based on three organizing principles identified in a detailed survey by Nacenta and colleagues [49, 50]: the referential domain, the display configuration, and the control paradigm.

Chapter Three describes the design and implementations of the auto-ordering techniques that were developed. Technique assessments of these solutions are also introduced in this chapter. At the end the chapter, we introduce some extended works based on our auto-ordering techniques.

Chapter Four details the comparative study that we conducted, which investigated the performance of our auto-ordering techniques in an object transfer task. Technique descriptions, study methods and results are reported.

Chapter Five explains the results found in the comparative study. The contributions and limitations of our auto-ordering techniques are also presented in this chapter.

Chapter Six presents a summary of the research of this thesis. Future directions are discussed.

CHAPTER 2

RELATED WORK

The chapter presents a survey of related research that form the foundation for the work presented in this thesis. We cover three main areas of research that serve as the background of our work: current solutions for determining devices' locations, F-formations and related topics in social organization, and approaches to transferring an object between mobile devices.

2.1 LOCATION SYSTEMS

A number of technologies have been explored to determine the locations of objects in mobile contexts. We divide these techniques into five categories according to the underlying sensor technology: infrared/ultrasound, Wi-Fi/Bluetooth beacons, motion capture, vision technology and sensors on mobile device.

2.1.1 Infrared/Ultrasound

Early systems used small transmitters to locate people and objects in an augmented environment. As early as 1992, Want et al. [67] developed a system, which detected the locations of person in an office environment, named Active Badge. People in this system carries an infrared badge that sends ID information to sensors located around a building through infrared signals. A central server then processes this information and distributes it.

To improve precision, some other systems use ultrasound to calculate position. Researchers at AT&T developed an ultrasound based positioning system, the Active Bat Location System [27]. In this system, users take a small sensor tag that emits an ultrasonic pulse to receivers mounted on the ceiling. The distance between Bat receivers are calculated by the times-of-flight of the ultrasound pulse form the Bat to receivers. When a Bat can "see" three or more receivers, the Bat's position then could be inferred using the process of multilateration [27]. Another ultrasound-based technique is the Cricket Location Support System [55]. Unlike Active Bat, the beacons mounted on wall and ceiling periodically transmit the ultrasonic pulse along with a RF signal which contains the location information. The listeners receive these RF and ultrasonic signals and perform all their own triangulation computations.

The Relate system [39] employs custom-built ultrasonic USB dongles to calculate relative positions to other devices. While the accuracy and precision were sufficient for the type of file transfer activities examined in this thesis, the custom hardware makes this solution unavailable to most users.

2.1.2 Wi-Fi / Bluetooth Beacons

Recent approaches have used existing infrastructure or active sensing to provide position information. For example, researchers have used trilateration with Wi-Fi signals to determine location. RADAR [5] is a positioning system that uses signal strength to locate and track objects. The basic approach used in RADAR is triangulation. The base station measures the signal strength sent by wireless devices, then determines the location that best matches the observer signal strength data. Similar in form to RADAR, some commercial wireless asset-tracking packages are available (e.g., the WhereNet real-time locating system [71]). Another Wi-Fi based positioning system SaskEPS [8] employs received signal strength indicators (RSSI) measurements, calibrated access point (AP) locations, and trilateration for indoor positioning. With the location information of the APs, this system can provide GPS-like accuracy [8].

The Virtual Compass system [6] uses Bluetooth received RSSI and Wi-Fi signals to calculate distances between devices and position them in a 2D plane. This system works without external infrastructure, but has a low accuracy – experiments showed that Bluetooth RSSI alone had a mean positioning error of 3.4m, that Wi-Fi alone had a mean error of 3.9m, and that a combined technique had a mean error of 1.4m, with error of 2.7m at the 90th percentile [6]. An error of 1.4m precludes the correct ordering of people standing around a circle. For example, taking a 1.4 m error and a two-meter diameter circular arrangement leads to an angular error of ±90 degrees along the circumference of the circle, making it implausible to use reliably to resolve ordering in F-formations.

2.1.3 Motion Capture

Motion-capture systems use magnetic tracking (e.g., the commercial Polhemus system [54]) or infrared cameras (e.g., the Vicon system [65]) to provide precise 3D positions, but only within the range of the cameras or antenna. Ascension Technology [3] provides a variety of motion-capture solutions. TrakSTAR is one of the product of their 3D Guidance family. This system uses external antennas and magnetic field transmitter to compute the position and orientation information.

Marquardt et al. [45] built a Proximity Toolkit to help developers easily obtain proxemics information in a room sized environment. Two motion tracking system are used in this research. The first one is a marker-based VICON [65] motion capturing system. Another is a Microsoft Kinect sensor. These two tracking systems are connected to the Proximity Toolkit Server which manages proxemics information. Developers can access this information through an event-driven programming library provided by the toolkit.

2.1.4 Vision Technology

The system developed by Wagner et al. [66] is built on two existing approaches: Scale Invariant Feature Transform (SIFT) and Ferns [53]. With this system, mobile phones can detect and track SIFT/Ferns features, calculate a camera pose and render augmented graphics at real-time frame rates.

Researchers have also investigated vision-based systems that track fiducial markers. Kato et al. [36] developed an Augmented Reality (AR) conferencing system. In their system, a user, who wearing a AR head mounted display (HMD), can see video images from remote desktop users. Remote desktop users can see the video image captured by the small camera on the HMD of the AR user. A set of markers on the AR user's side help the corresponding remote users on the HMD. When running the system, computer vision techniques can help identify and draw remote users on their marker counterpart.

Kray et al. [40, 41] used external cameras to detect markers on devices. They proposed an interaction technique based on dynamically generated spatial regions around mobile devices, for a photo sharing task. In their method, an external camera and PC are used for detecting the markers shown on the screens of mobile devices. According to different radii, they defined three circular spatial proximity regions with the center of a mobile device. Other mobile devices in different regions would trigger different behaviors, such as image preview or photo download.

Another AR based system developed by Li et al. [42] used the front facing camera of mobile devices to combine multiple devices together as an interactive display surface. They use the front facing camera to track a fiducial marker on the ceiling. Schwarz et al. [59] use a color-transition encoding scheme to identify and locate displays. First, each device in the group fetch a unique ID (color sequence) from a web server, then, these devices display the appropriate color on their screens simultaneously. Meanwhile, a camera records these colors and send them to the web server. After each device finished displaying its color sequence. The server resolves the devices' relative locations and assigns display content to them.

Dearman et al. [17] present a method to determine the relative orientation of proximate devices using device's backside camera. Computer vision techniques are used to extract common features from images and then compute the orientation. They use GPS [22] or Wi-Fi position, which are integrated in the Android OS, to determine which client devices are proximate. A web-service is used to do the heavy work – orientation calculation.

2.1.5 On-device Sensors

Chen et al. [14] used the phone's built-in sensors to detect the spatial relationship between user and mobile device. Four built-in sensors are used in this study. The relative distance between the device and user's face is calculated by the head size of the user from the front camera image. Orientation sensors (gyroscope, accelerometer and magnetometer) are used to track the device's horizontal orientation and vertical orientation.

SurfaceLink [23] is a system which combines accelerometers, vibration motors, speakers, and microphones, to detect device positions on a shared surface. The basic idea is that devices on the same surface could sense the same vibration patterns and then detect the relative arrangement.

Jin et al. [34] present Tracko that synthesized Bluetooth low energy signals and inaudible stereo sound to deduce 3D locations of nearby devices. The Bluetooth low energy signal is used to detect the presence of other devices. They use the arrival times of exchanged inaudible stereo signals to estimate the distance and 3D direction between devices.

2.2 SOCIAL ARRANGEMENT

2.2.1 The F-formation System

Researchers have studied the ways in which people organize themselves when they come together as a group (e.g., [20, 37, 38, 46]). In particular, the physical arrangements that people use have been examined by Kendon [37, 38], who determined that people often arrange themselves in roughly circular clusters called Facing Formations (F-formations).

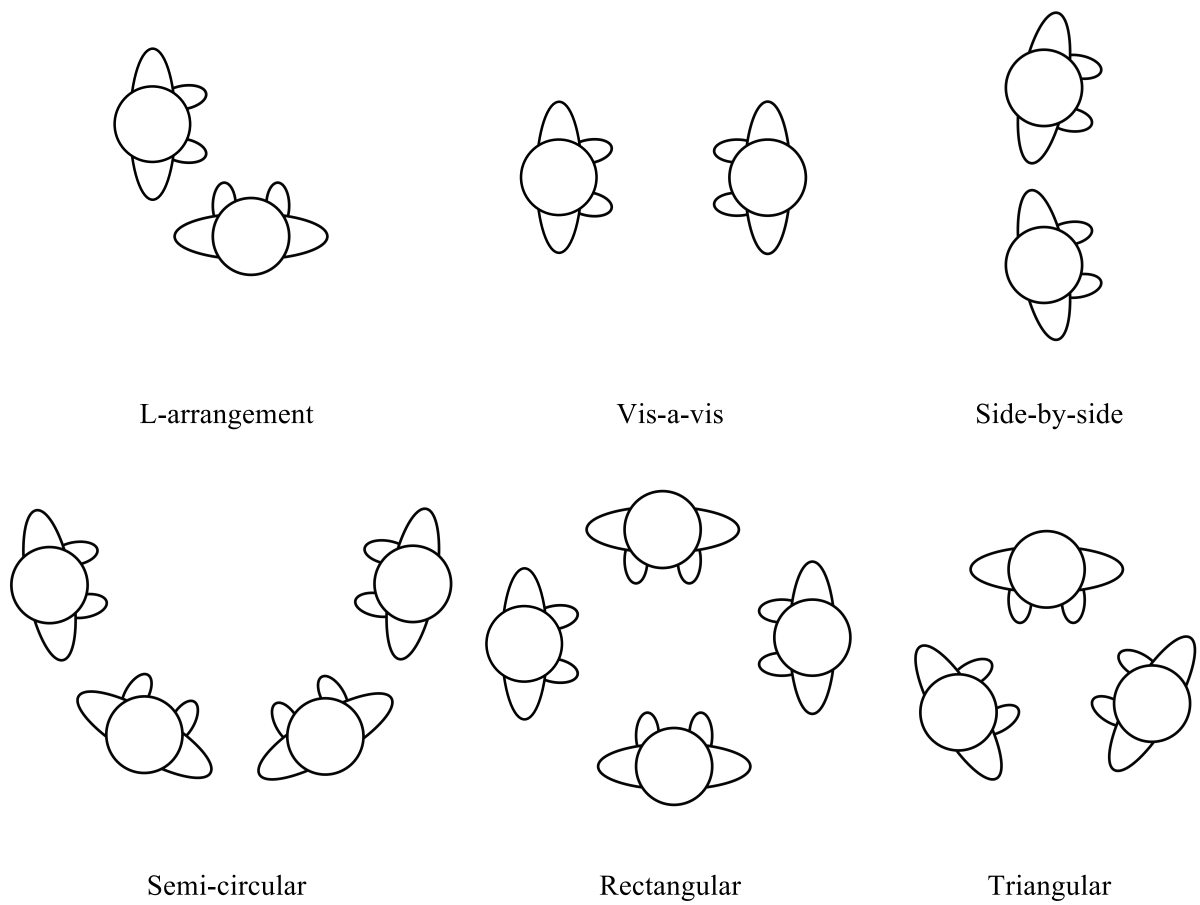


Figure 2.1: Some F-formation arrangements

Quoting Kendon, “an F-formation system arises when two or more people cooperate together to maintain a space between them to which they all have direct and exclusive access.” [37]. “F-formations are characteristic of people who come together to accomplish joint activity.” [15]. F-Formations can occur in many different settings, may be physically larger or smaller depending on the situation (e.g., a conversational group may have a smaller arrangement than a work group around a table), and may be only approximately circular (e.g., L-arrangement, vis-a-vis, side-by-side, rectangular or triangular arrangements are also possible). Figure 2.1 shows some F-formation arrangements. F-Formations typically comprise between two and five people [46], and gestures or objects within the space between these people can become the focus of the interaction. We can easily find these arrangements in our daily social interactions (see figure 2.2).

Figure 2.2: F-formations in social interaction.[1]

a. face-to-face; b. side-by-side; c. triangular

The arrangement adopted by an F-formation depends on three factors [37]. The first is the number of participants. For example, a side-by-side arrangement is unlikely to be formed with more than three participants. Secondly, the environment where the activity happened also determine the shape of F-formation. A face-to-face arrangement commonly occurs when two people are talking. When a conversation happens in front of exhibition display screens, a side-by-side arrangement is

[1] Author: Wikimedia Israel Source: Wikimania 2011 Pre-Conference
https://commons.wikimedia.org/wiki/File:Flickr_-_Wikimedia_Israel_-_Wikimania_2011_Pre-Conference_(75).jpg

more likely to be formed. Third, the social activity performed by participants can also affect the F-formation arrangement. For instance, a group meeting can occur with a circular arrangement. While, a meeting held by a manager who is leading the conversation commonly happens in the form of L-arrangement [37]. Ciolek et al. [15] investigated the different F-formations adopt by two people in four dissimilar contexts. Their result showed that circumstances did impact the frequency of occurrence of various types of spatial arrangements [15].

Kendon [37, 38] described three kinds of spaces to delineate the domains of an F-formation (figure 2.3). He first defined a concept named a transactional segment. “It is the space into which a person looks and speaks, into which he reaches to handle object” [37]. When several people gathered together to do the same thing, their transactional segments overlap. The term *o-space* is used to delineate this joint transactional space. The *o-space* is the core territory where the main activity the group is pursuing occurs [46]. The second spatial domain is *p-space* which is a narrow strip around the *o-space*. The bodies and personal belongings of people involved in F-formation are placed in this domain [15]. Beyond the *p-space*, a further area is called *r-space*. This space excludes the group from the outside world. It also works as a buffer zone where the people who might join or leave the system can be positioned [46].

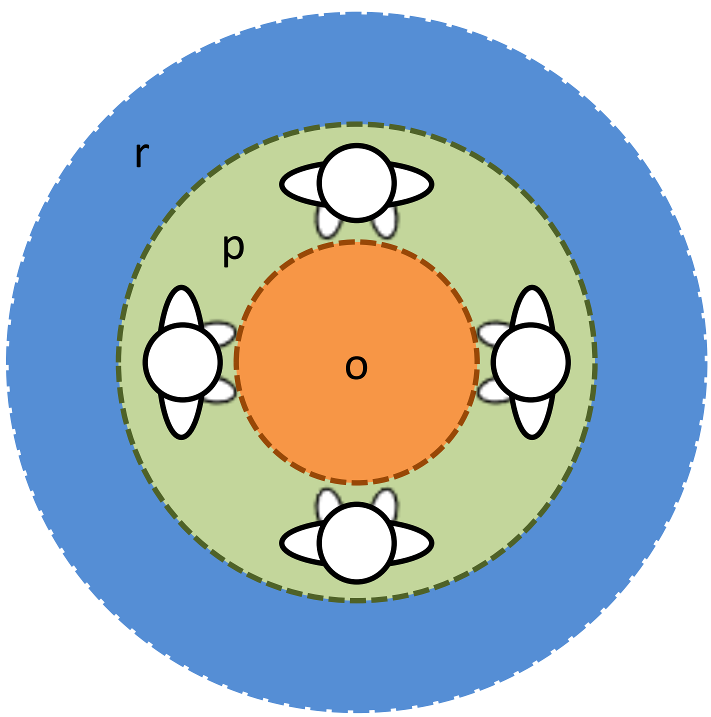


Figure 2.3: The domains of an F-formation

2.2.2 F-formations in HCI

Researchers have started to model groups of interacting people. Marshall et al. [47] used the concept of the F-formation system to analyze the social interactions in a tourist information center. Cristani et al. [16] investigated detection of social interactions using the estimated people's positions and head orientations to detect the *o-space* based on a Hough-voting strategy. Taking the result of the estimation of the *o-space*, they could identify the people who form the group. They then could recognize people who are socially interacting.

Yu and colleagues [70] applied face recognition algorithms on video capture from surveillance system to discover social groups. Hung and Kröse [33] improved on this work by analyzing social interactions in crowded environments. Similar to [16], people's positions and head orientations are employed, but F-formations were identified based on a graph-theoretic clustering algorithm. Other F-formation detection methods using graph-theoretic clustering algorithms are described in [62, 63] and [64]. Tran et al. [62, 63] proposed a graph-based clustering algorithm to discover interacting groups in crowded scenes. Social signaling cues are used to generate a graph that represents a person and their interactions. Vascon et al. [64] developed an approach which combined the modeling of the uncertainty in the position and orientation of a person and a game-theoretic clustering approach.

A comparative study conducted in [60] showed that the approach described in [16] worked better than the one described in [33] when position and orientation are known. When only position information was available, Hung's algorithm had better performance. In 2015, Setti et al. [61] introduced a novel methodology called Graph-Cuts for F-formation (GCFF). Based on a graph-cuts framework for clustering individuals, GCFF could automatically detect groups in still images [61].

For the purpose of automatically taking photos with a set of preset cameras, Gan et al. [21] used Kinect depth sensors to generate heat map which represents the spatial location, orientation, and temporal information in F-formation systems.

HCI researchers also have used F-Formations as the basis for interaction techniques. Marquardt and colleagues [46] developed techniques that allow easy object transfer when people are beside one another, that provide screen previews based on device tilt, and that allow full screen sharing when people are in an F-Formation. In this system, Kinect cameras, mounted on the ceiling, are used to track small groups of people. In addition, low-power 8GHz band radio modules and accelerometers are used to detect devices' locations. However, these techniques required an external tracking system making them difficult to use in truly ad-hoc interactions.

2.3 OBJECT TRANSFER TECHNIQUES

Several researchers in HCI have considered the problem of how to move objects from one device to another in multi-display environments. Below, we summarize this research based on three organizing principles identified in a detailed survey by Nacenta and colleagues [50]: the referential domain, the display configuration, and the control paradigm (also see Figure 2.4). Other researchers such as Rädle et al. [57] have proposed extensions to this architecture.

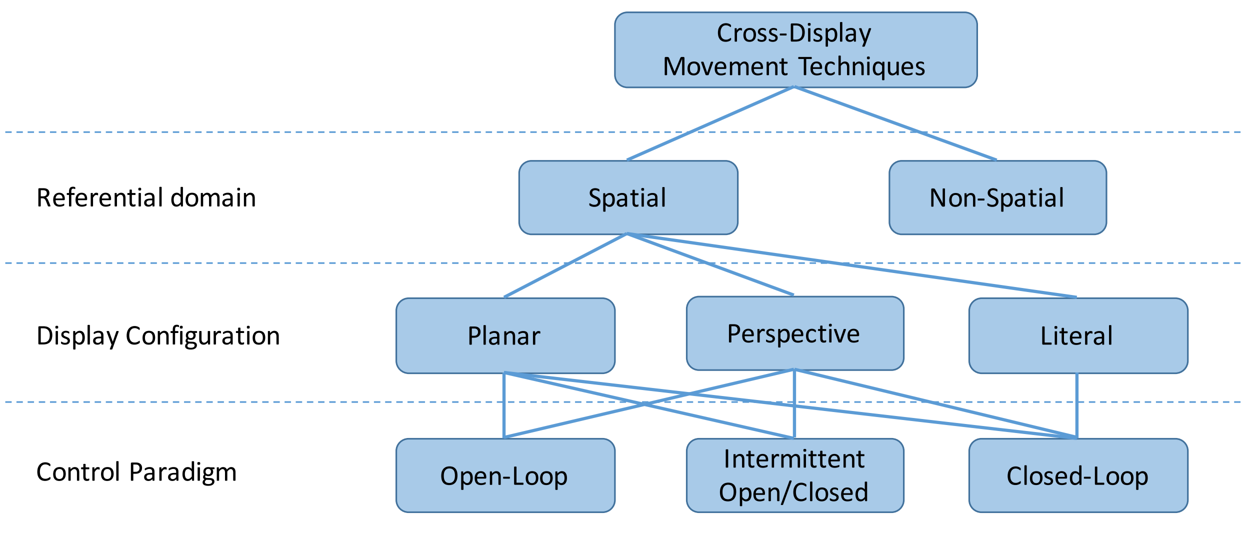


Figure 2.4: Graphical representation of Nacenta's three organizing principles [50]

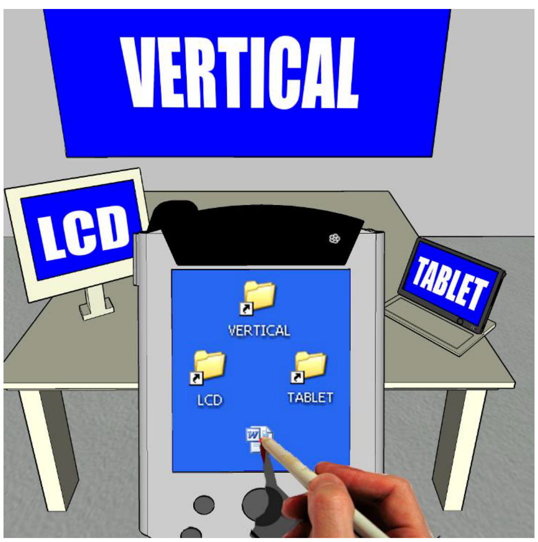
2.3.1 Referential Domain

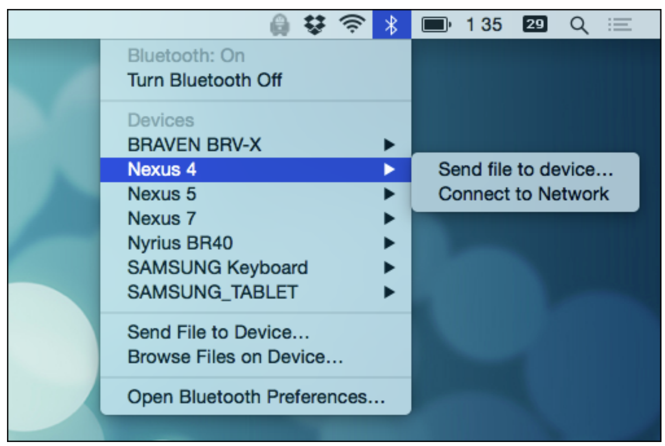
The referential domain is the way in which users refer to different displays. Two main reference types that shows in Figure 2.5 are spatial arrangement (e.g., the display locate on the right) and named displays (e.g., the display named "Leslie's monitor"). Correspondingly, a cross-display movement technique might require a direction or a name of the target display as an input.

Several object transfer techniques have been developed for both types: for example, real-world spatial locations were used with the early techniques. The Put-that-there technique [12] enables users to move objects in a large-screen graphic display by voice and pointing gesture. In Pick-and-

Drop [58], users pick up an object on a display and drop it on another display by a digital pen as if they are moving object in real world. On-screen representations of real-world locations are used with “world-in-miniature” views (e.g., ARIS [10] enables users to relocate applications to remote display via a mini-map that shows the physical arrangements of displays), and arbitrary spatial locations are used with many portal-based techniques.

Named displays, in contrast, use non-spatial methods such as text, numbers, or colors to refer to other devices. Many techniques have used this approach. Multibrowsing [35] is a framework allows users to move web content among multiple displays by choosing the target device from a menu shows in a popup window. Mighty Mouse [13] maps other displays with a list of names or icons. Conductor [25] uses color-coded icons to represent different displays. Contact lists and shared-folder icons are also popular in commercial systems such as Bluetooth [11] allows users to select targets from a list of connected devices. Finally, some techniques allow users to cycle through a set of displays (e.g., Multi-Monitor Mouse [9]) by pressing a key or button, as if switching application in Windows system by shortcut keys.





(A) spatial [49]

(B) non-spatial

Figure 2.5: Two cross-display object movement interaction techniques

2.3.2 Display Configuration

The display configuration is the way in which the displays are organized in physical and digital space. This dimension affects techniques that use direct manipulation to transfer objects, because the arrangement of display limits the kinds of transfer actions that can occur. Nacenta categories existing techniques according to their input models into three groups: planar, perspective and literal.

Some techniques have arranged displays in a "stitched" fashion, where the different screens form a single workspace. Stitched displays allow object transfer by moving the object across the edge of the display (e.g., [46]). A typical stitched display configuration is the multi-monitor mode in current operating systems [50] (see figure 2.6). A stitching can cause problems when different users see the displays from different directions [50], and so other techniques use the perspective of the user to organize display locations (e.g., the Perspective Cursor system [51]). Finally, "literal" techniques can use the actual devices themselves rather than their locations to enable object transfer. Stitching [31] is an interaction technique uses synchronous gestures to combine multiple tablets. Holmquist et al. [32] developed an interface called "Smart-Its Friends" which establishes connection between two devices by holding them together and shaking them. The Android's Near-Field Communication technique connects two devices when they are physical proximity. Hinckley explores bumping [30] as the mechanism for indicating which device is the target. One main drawback of these literal techniques is that they are limited by the physical reach of the user [50].

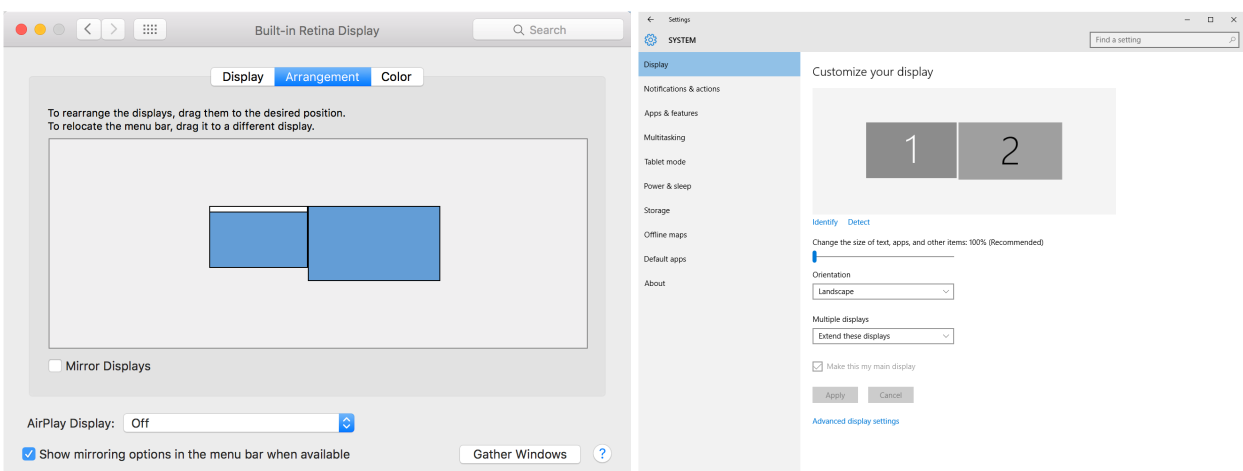


Figure 2.6: Two stitched display configurations in OSX and Windows operation systems

2.3.3 Control Paradigm

The control paradigm is the way in which people actually perform the transfer. According to Nacenta's theory, there are three control possibilities for cross-display interaction techniques: open loop, closed loop and intermittent open/closed control.

Open-loop transfer does not provide feedback to users, such as the Flick [68] and Multi-Monitor Mouse [9] systems. In contrast, Closed-loop transfer enables users to control the action during the whole transfer process. This kind of control type is used by all techniques that have a visible representation of the target (including world-in-miniature systems [12], pantograph-style movement [28], and portal-based techniques [25]). Some other techniques (e.g., Mouse Ether [7] and Perspective Cursor system [51]) account for the blank space between displays in the physical world. These techniques belong to the intermittent control type since when the cursor is in displayable space, the process is closed-loop, when it is in blank space, the process is open-loop [50].

In this chapter, we reviewed previous researches on determining devices' locations. The lack of a lightweight positioning technology forms the initial idea of this thesis. As our theoretical basis, F-formation system enables us narrow the problem from detecting precise locations in a plane to calculating relative locations about a circle. We reviewed the three organizing principles for object transfer technique in multi-display environments. In terms of these three dimensions, the auto-ordering techniques we developed use a spatial referential domain (using real-world locations), a perspective-based display organization (i.e., targets are arranged correctly for each person's view), and either open-loop or closed-loop control (since the technique supports both flicking and portal-based transfer).

CHAPTER 3

AUTO-ORDERING TECHNIQUES

People frequently form small groups in many social and professional situations. Using the relative position in these groups for file sharing could facilitate intuitive interfaces. In this research, we developed three systems (marker-based, compass-based, and light-gradient-based ordering) that can auto-order locations about a circle based on sensors standard on commodity smartphones. In this chapter, we will introduce the overall architecture and the three auto-ordering techniques. Furthermore, we have made some extensions to our system which will be presented later in this chapter.

3.1 OVERVIEW

Auto-ordering of people engaged in F-Formations can be viewed as the technical problem of determining the relative location of the users, and faithfully rendering the relative locations on each user's device. The general problem of determining relative location can be quite complex, as it requires determining the position and orientation of individuals with respect to a common coordinate frame. General positioning technology using GPS, Wi-Fi, or Bluetooth localization do not have the spatial fidelity to resolve the relative locations of individuals standing in a typical F-formation, and dedicated hardware with better spatial fidelity like IR tracking systems can be cumbersome or difficult to install.

Our assumption of users in an F-Formation allows us to constrain the problem to the point where sensors commonly found on mobile devices can perform auto-ordering registration.

Figure 3.1 shows a schematic representation of the problem. Given four individuals (three shown, one holding the phone), we need to share a file with only one of them in an ad-hoc network. Labeling can happen through tags (e.g., color) or through relative location (ordering around the circle). If the ordering were arbitrary, for example the purple and orange recipients were flipped, it would be more cognitively difficult to assign the appropriate tag to the appropriate on screen location.

Our design goals are to provide rapid operation that facilitates sharing, to minimize user error, and to have little or no physical setup required. We have developed three solutions, employing different sensor suites common to today's smartphones. The first is a marker-based technique that uses a fiducial marker to provide a visual reference to which each phone can calculate its relative position. The second technique leverages the orientation sensors on the phone (accelerometer, gyroscope and compass) to determine the relative orientation of each user, which is then mapped to a circle around which they are standing. The third technique uses the phone's camera to determine a direction from the background between the devices.

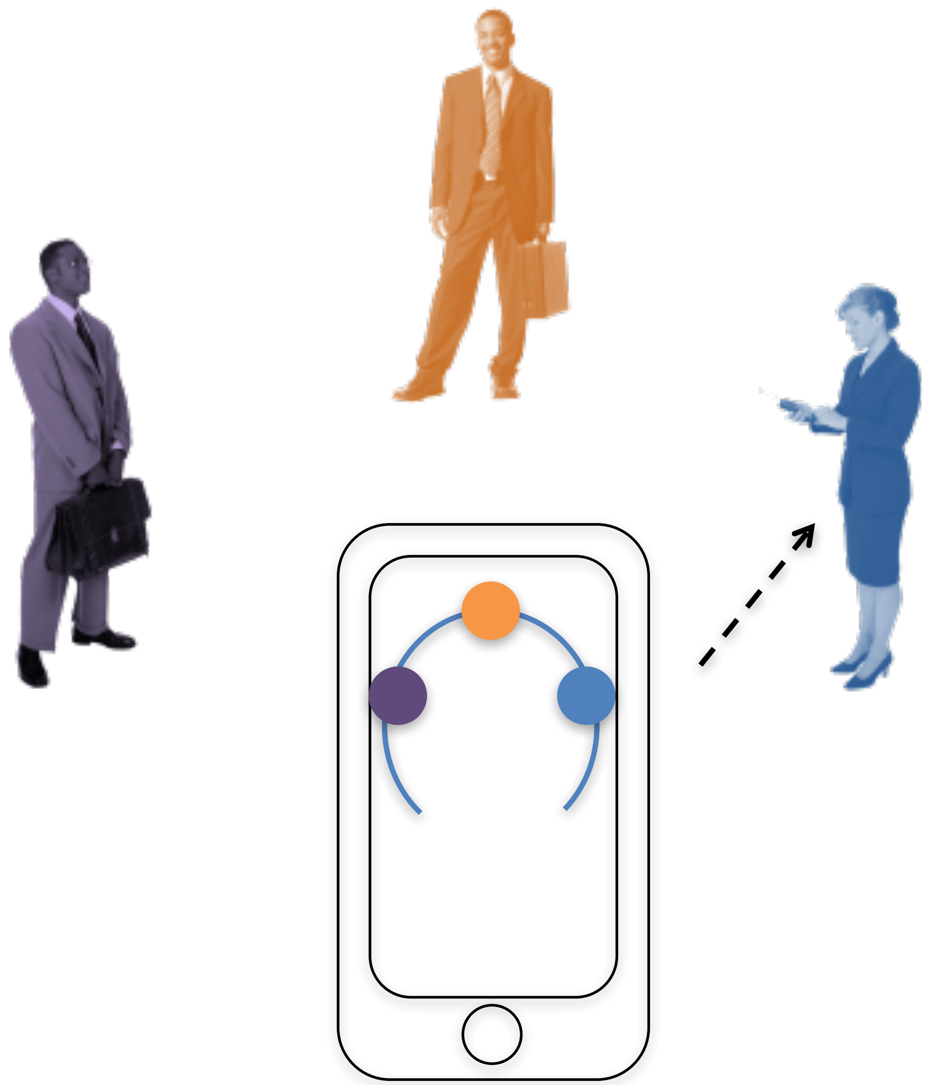


Figure 3.1: A typical transfer setting: four people in a circle, and the person holding the device must determine which on-screen portal corresponds to which person in the real world

We assume that users roughly face each other (i.e., they form simple open or closed shapes where people face inwards) and are at roughly the same distance from the group's center, which is common in self-ordering behavior of small groups [38, 46]. These assumptions provide us with a critical insight that allows us to address the problem more accurately and reliably than previous attempts. Because we can assume that users are arranged around a circle, we can constrain the solution space to that manifold. To express ordering about a circle, one does not need relative location, but rather relative orientation – the polar coordinates of the person on the circle. Because the radius of the circle is fixed, or at least quasi-static with respect to the interactions, only the angular coordinate is required to determine relative position.

A key component of auto-ordering techniques is that they allow portal locations or flicking directions to be arranged to match the location of the actual people or devices in the real world. The general psychological principle of stimulus-response compatibility [56] predicts that digital arrangements that correspond to the physical world will be faster and produce less cognitive load because they allow people to use information provided by the real world instead of having to remember an arbitrary mapping (e.g., it may be easier to flick a document toward a real-world printer than selecting the printer from a list). Nacenta [49] showed mixed results when applying the idea of stimulus-response compatibility to transfer tasks – and no published experiments have assessed the use of world-to-interface correspondence for object transfer.

3.2 ARCHITECTURE

Our auto-ordering systems are built on Android platform, an open source project initiated by Google, Inc. Different techniques leverage different sensors within the Android library. As shown in Figure 3.2, the marker-based technique and the light-gradient-based technique depend on third party libraries, AndAR [18] a variation of ARToolkit [36], and OpenCV [52] respectively. The Compass-based technique relies on Android's sensor manager (accelerometer and magnetic field). Bluetooth is used for the data transport layer to facilitate exchanging location information.

In our application design, a Bluetooth server is running on one of the phones in the group. Mobile devices connected to the same server are considered in the same group. Each device sends its location information to the server on every system heartbeat (100 milliseconds in our design). The Bluetooth server broadcasts all devices' location periodically.

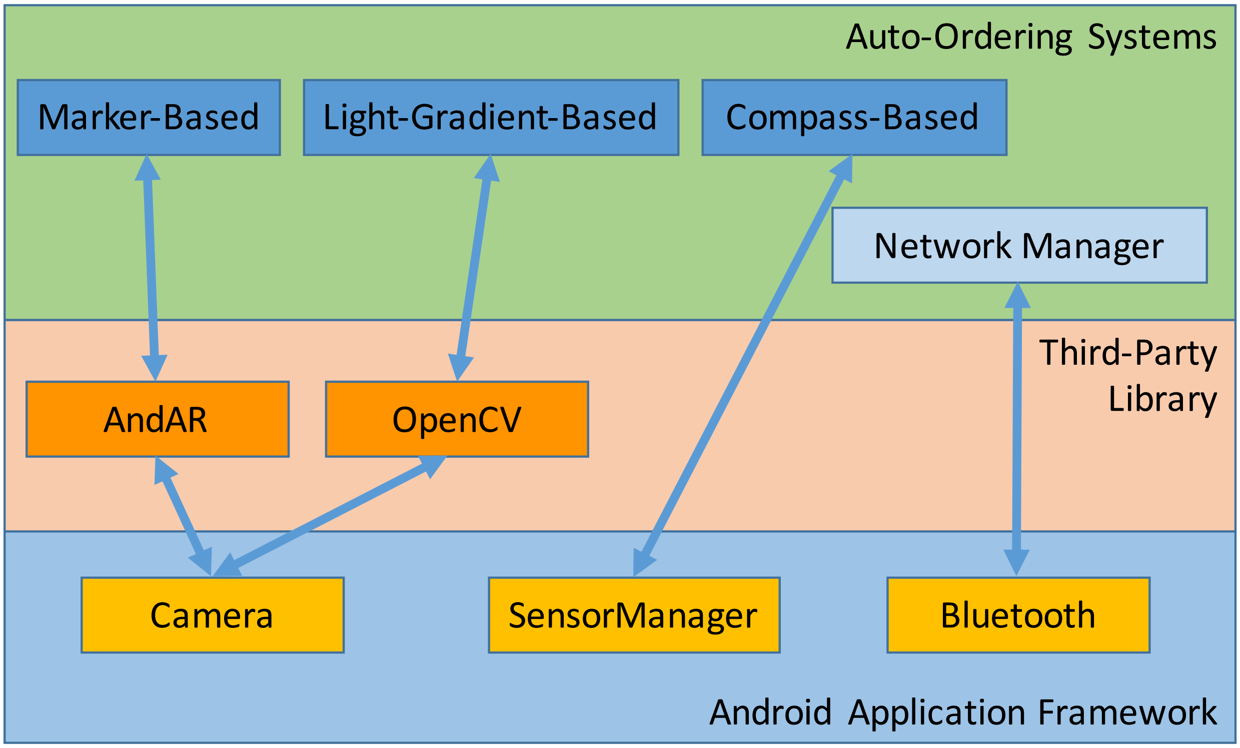


Figure 3.2: Auto-ordering systems architecture

3.3 MARKER-BASED ORDERING

A visual or fiducial marker is a standardized shape or mark usually a heavy square, easily detected using image processing techniques. Examples of visual markers are shown in Figure 3.3. Once users have organized themselves in an F-Formation, it is trivial to introduce a marker at the center of the circle, either as a piece of paper, or more likely as the screen of one of the participant phones. This section will describe the marker-based auto-ordering technique and its related technology.

Bar Code QR Code AR Tag

Figure 3.3: Visual Markers

3.3.1 Augmented Reality

3.3.1.1 Definition

In general word, Augmented Reality (AR) mixes virtual object into the real world. According to the research of Azuma [4], an AR system must have the following characteristics:

1) Combines real and virtual
2) Interactive in real time
3) Registered in 3-D

3.3.1.2 ARToolKit

ARToolkit [2] is a software library used for building Augmented Reality applications, which uses orientation sensors and image processing to detect the pose of a phone with respect to a marker, and optionally can render virtual objects over the marker. The shape of an ARToolkit marker can theoretically be any image as a marker pattern, surrounded by a black square [36], as shown in Figure 3.4.

Figure 3.4: ARToolKit Marker

The library provides the following features [2]:

- Single camera position/orientation tracking.
- Tracking code that uses simple black squares.
- The ability to use any square marker patterns.
- Easy camera calibration code.
- Fast enough for real time AR applications.
- SGI IRIX, Linux, MacOS and Windows OS distributions.
- Distributed with complete source code.

3.3.1.3 AndAR

ARToolKit is originally written in C code. It is design for desktop development. For the purpose of running an AR library on smart phones, we employed the AndAR [18] library, which is an open source variant of ARToolkit for Android. This library is built upon the native ARToolKit library and Android library. It offers JAVA APIs, which makes the detection of ARToolKit markers on mobile device easily accessible.

3.3.2 Relative Orientation

AndAR library returns a matrix representing the pose of the phone with respect to the marker of the form:

$$\begin{pmatrix} r_{11} & r_{12} & r_{13} & t_x \\ r_{21} & r_{22} & r_{23} & t_y \\ r_{31} & r_{32} & r_{33} & t_z \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

The sub-matrix ($r11 \ldots r33$) represents the linear transformation, scale, rotation and shear of the object. The values of the fourth column (tx, ty, tz) contain the coordinates of the marker with respect to the virtual camera. We used the sub-matrix ($r11 \ldots r33$) to calculate the relative rotation between marker and phone. The center of the circle is defined by the fiducial marker. Once each phone has its angle encoded as ($r11 \ldots r33$) it is trivial to resolve the position around the circle. Users on a display can then be rendered with respect to the user with angular differences.

When people group together in the form of an F-formation, as in Figure 3.5, four individuals arrange as a circle shape around the marker, they can naturally define the center of the circle with the fiducial marker. Each phone has its angle with respect to the marker. Users know they are in the same group rendered by the color on the device screen.

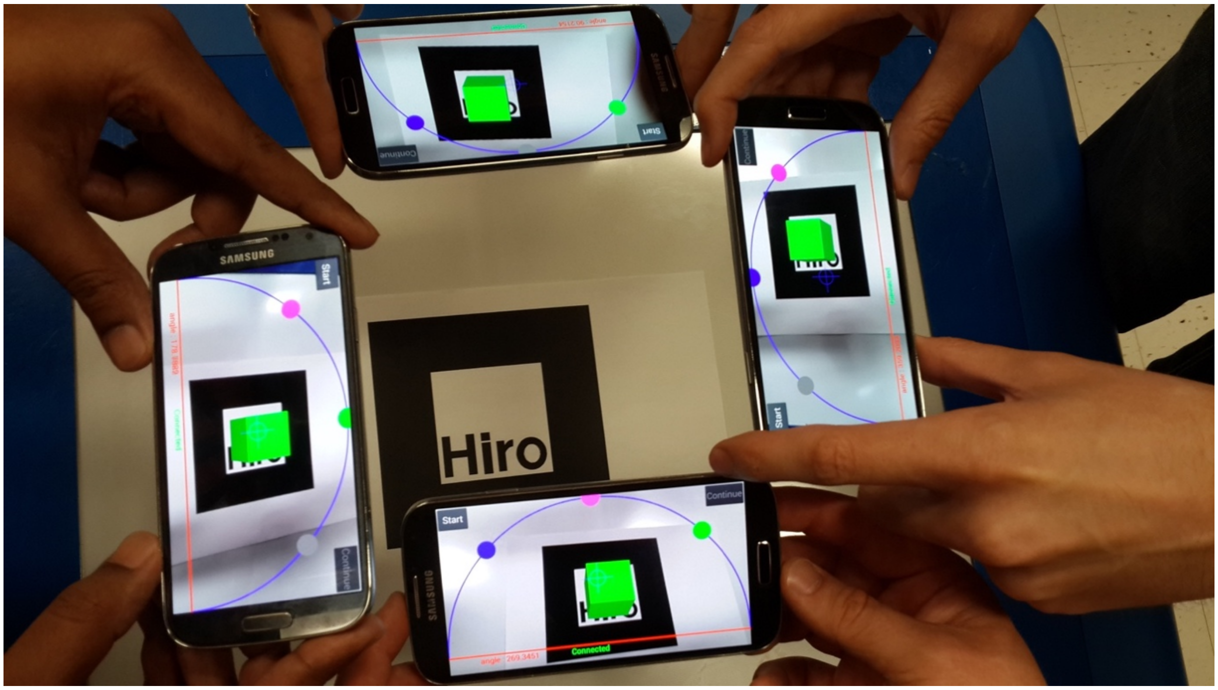


Figure 3.5: Marker-based auto-ordering system

Typical implementations employing QR codes use paper markers. However, the need for a dedicated marker, could be inconvenient in many of the scenarios described earlier. Therefore, we extended the technique to use a marker that is displayed on one of the phones in the circle (see Figure 3.6). In this variant of the technique, one person moves their device forward so that it is in view of the other phone's cameras, and this central phone displays a marker that is similar to the paper version.

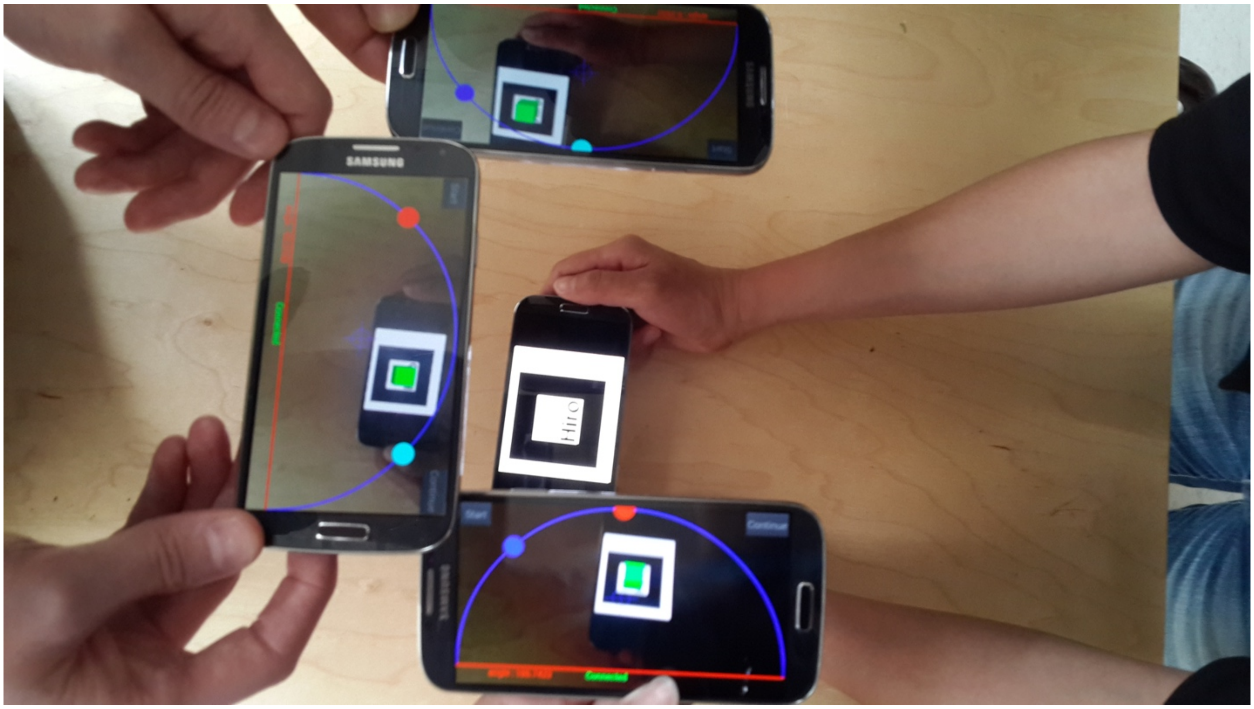


Figure 3.6: Marker technique using one of the phones to display the marker

3.4 COMPASS-BASED ORDERING

Most mobile devices provide sensors which detect the orientation of the phone with respect to the Earth, with the direction of gravity providing the vertical axis and a compass heading providing orientation around that axis. Compass heading can be translated into relative position about a circle if the pose of the phone with respect to the circle is fixed.

3.4.1 Compass on Android

Fusing data from the compass, gyroscope and accelerometers, Android provides an abstract sensor class Orientation describing the phone's pose. Figure 3.7 shows the Coordinate System (relative to a device) that's used by the Sensor API. According to Android's development documentation [1], X is tangential to the ground at the device's current location and points roughly East. Y is tangential to the ground at the device's current location and points toward the geomagnetic North Pole. Z points toward the sky and is perpendicular to the ground plane.

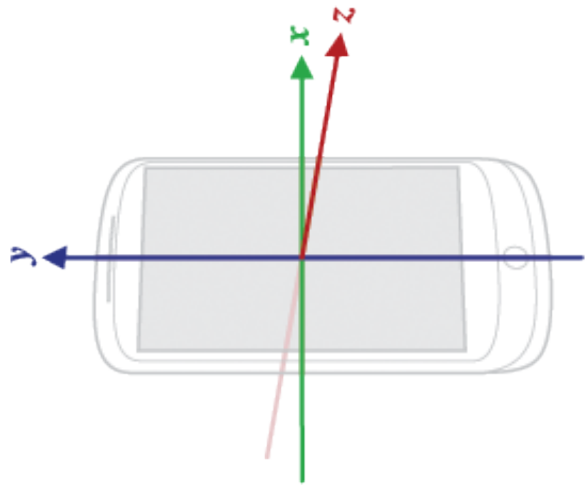


Figure 3.7: Coordinate System (relative to a device) that's used by the Sensor API

The Android SDK provides a function SensorManager.getOrientation() to calculate device's orientation. Before applying this function, Android must first determine the pose of the phone with respect to the north shown in figure 3.8. Once we have the rotation matrix, we can compute the device's orientation based on it via SensorManager.getOrientation(). This function gives us an array that contains the device's orientation data.

As introduced in [1]:

"values 0: Azimuth, angle of rotation about the -z axis. This value represents the angle between the device's y axis and the magnetic north pole. When facing north, this angle is 0, when facing south, this angle is π. Likewise, when facing east, this angle is π/2, and when facing west, this angle is -π/2. The range of values is -π to π.

value 1: Pitch, angle of rotation about the x axis. This value represents the angle between a plane parallel to the device's screen and a plane parallel to the ground. Assuming that the bottom edge of the device faces the user and that the screen is face-up, tilting the top edge of the device toward the ground creates a positive pitch angle. The range of values is -π to π.

value 2: Roll, angle of rotation about the y axis. This value represents the angle between a plane perpendicular to the device's screen and a plane perpendicular to the ground. Assuming that the bottom edge of the device faces the user and that the screen is face-up, tilting the left edge of the device toward the ground creates a positive roll angle. The range of values is -π/2 to π/2."

3.4.2 Smooth Data

To reduce data jitters, we applied two low-pass filters to the compass readings.

The first low-pass filter passes low-frequency signals but attenuates signals with frequencies higher than the cut-off frequency. It provides a smoother signal, removing the short-term fluctuations. Figure 3.8 shows the implementation of the low-pass filter. The APLAH value showed in the code is the time smoothing constant for low-pass-filter. The smaller the value the more smoothing applied to the filter.

```java
public class LowPassFilter {
   /*
    * Time smoothing constant for low-pass filter 0 ≤ α ≤ 1 ; a smaller value
    * basically means more smoothing
    */
   private static final float ALPHA = 0.2f;

   private LowPassFilter() {
   }

   /**
    * Filter the given input against the previous values and return a low-pass
    * filtered result.
    *
    * @param input
    *          float array to smooth.
    * @param prev
    *          float array representing the previous values.
    * @return float array smoothed with a low-pass filter.
    */
   public static float[] filter(float[] input, float[] prev) {
      if (input == null || prev == null)
         throw new NullPointerException("input and prev float arrays must be non-NULL");

      if (input.length != prev.length)
         throw new IllegalArgumentException("input and prev must be the same length");

      for (int i = 0; i < input.length; i++) {
         prev[i] = prev[i] + ALPHA * (input[i] - prev[i]);
      }

      return prev;
   }
}
```

Figure 3.8: Low-pass filter

The second low-pass filter algorithm used in our compass-based auto-ordering system is a moving average filter on the top of a data buffer cache (shown in Figure 3.9). We created a buffer that caches the most recent thirty compass data points. A first in first out (FIFO) queue is the underlying data structure. Every updated data passed into our auto-ordering system is the average number of the thirty data in the queue, implementing a moving average filter.

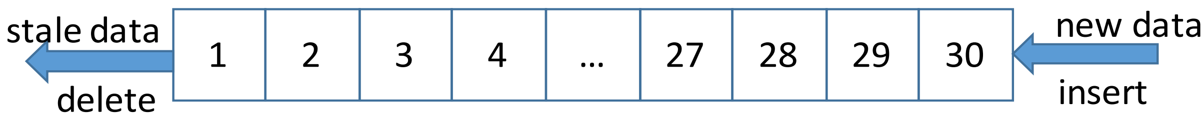


Figure 3.9: FIFO Queue

3.4.3 Relative Orientation

In the compass-based registration system, users point their phones at the arbitrary center of a circle. If users are arranged in a closed shape such as a circle or square, this is trivial, as the arbitrary center is in the center of the shape they form. For more truncated shapes such as line segments or semi-circles, the alignment is only slightly more complex, and can be aided by selecting objects in the real world to serve as the approximate center of the circle.

Once all users are pointed at the center of the circle, their positions around the circle can be inferred by the relative orientation they have with magnetic North. Assuming that magnetic North is always zero, then the relative location of each participant around the circle is simply the angular location of their reported orientation minus the angular location of the current user.

As shown in Figure 3.10, four mobile phones are pointing to the center of a circle shape. Mobile devices in the same group are represented by colored circles. For example, the bottom right phone is displayed as a blue circle on the screens of the other phones in the group.

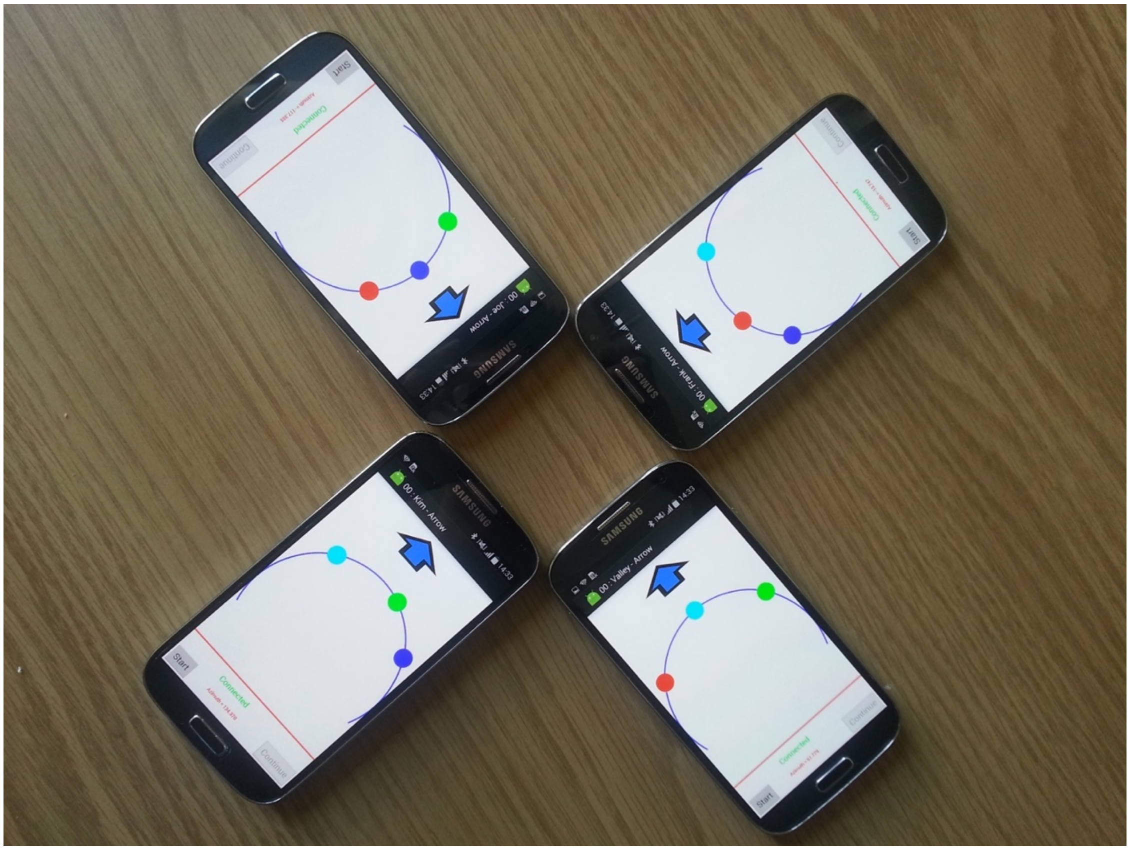


Figure 3.10: Compass-based auto-ordering system

3.5 LIGHT-GRADIENT-BASED ORDERING

Several researchers have investigated the automatic registration of images taken from different perspectives – for example, Microsoft's PhotoSynth (photosynth.net). These systems examine a set of images to look for features that can be matched with the other views, allowing the system to calculate relative device poses. Inspired by this idea, we developed a light-gradient-based auto-ordering system based on the OpenCV [52] library.

3.5.1 OpenCV

The OpenCV (Open Source Computer Vision) [52] is a free cross-platform real-time computer vision library distributed under the BSD open source license. It contains a number of image-processing functions and algorithms such as face detection, pedestrian detection, feature matching, and tracking.

The OpenCV library is originally written in C and C++. It also provides JAVA interface for Android development. In this thesis, we use the OpenCV for Android library (version 3.0).

3.5.2 Image Processing

Our light-gradient-based auto-ordering system uses the OpenCV library to perform the process of image capturing and processing. The OpenCV library offers a suitable interface for users to create a camera view and updates the view on each frame. The process flow shows in Figure 3.12 describes how our system works.

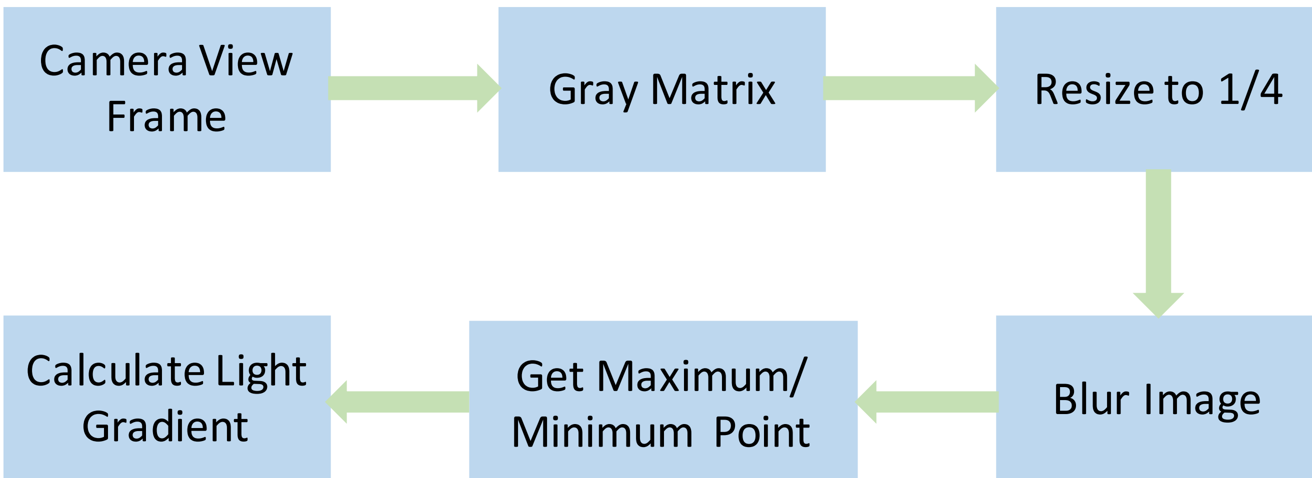


Figure 3.12: Light-gradient image process

(1) Read image data from camera view
(2) Get gray matrix from the image data
(3) Resize image to ¼ of the original size by bilinear interpolation
(4) Smooth the image using a 3x3 normalized box filter
(5) Get the brightest and darkest point in the image
(6) Calculate light gradient direction via the location of the brightest and darkest point

3.5.3 Light-Gradient-Based Ordering System

We used a simple approach to allow mobile phones to detect relative ordering around a circle. As described in section 3.5.2, we use the brightness gradient of the table or floor between the phones as the single feature to be identified. Each phone determines a direction by drawing a line between the darkest and brightest points on the background, and then communicates this direction to the other phones. Because they all see the same gradient but from different angles, they can determine their relative ordering by comparing the reported directions from each phone.

The system is shown in Figure 3.13. The current version uses a simplistic feature-recognition approach which works best when strong lighting gradients are evident, for example from spotlights or through open windows, but works poorly under diffuse lighting. Our implementation demonstrates that the idea is possible, but more robust feature-detection algorithms are required before the technique achieves the same degree of generalizability as the compass technique. Although the gradient based technique performed well in controlled laboratory conditions, its lack of generalizability in its current form lead us to remove it from consideration in the usability study.

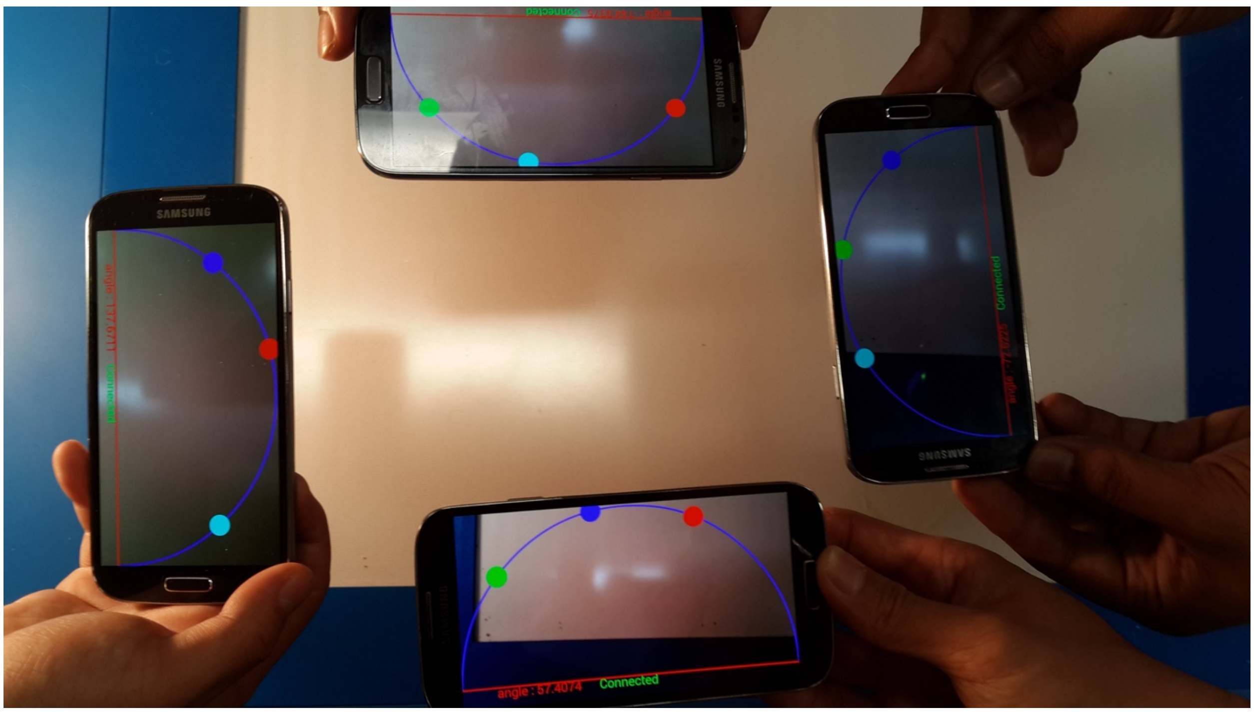


Figure 3.13: Light-gradient-based auto-ordering system, using the lighting gradient on a tabletop

3.6 TECHNICAL ASSESSMENT

For the proposed systems to be functionally useful, the sensors must return sufficient angular resolution to localize a sufficient number of users in a circular arrangement. We consider three classic parameters: sensitivity, (the degree to which sensed values represent reality), precision (the degree to which a sensor returns the same value for the same stimulus), and span (the range over which sensor readings are valid), to characterize the sensors and to determine the number of people that can be reliably localized. We hold that at least five people should be reliably localizable around a semicircle, based on Bahl's work [5], therefore an angular resolution of at worst 36° is required. We note that the Bluetooth-based Virtual Compass [6] is insufficient for practical use, as its mean angular error is almost triple this requirement.

The device used for testing was the Samsung Galaxy S4 (1.6 GHz processor, 5-inch 1080p display, Android 4.4.2).

3.6.1 Marker-Based Ordering

Fiducial markers can provide highly accurate (cm-scale) and stable pose estimates from camera images, ensuring that ordering will always be maintained. Because of the high-fidelity cameras on modern smartphones, the accuracy and precision of the marker-based localization are well below the 36° threshold specified. Typically, angular positions in of less than 3 degrees could be easily resolved. The primary limiting factor in the marker case is not the accuracy or precision, but the span, because the marker must always remain in view of the camera, constraining the number of possible angles. Two parameters of span must be considered: the distance and angle at which the marker can be resolved.

To test the maximum detectable distance, the markers were mounted on a wall. Marker size in pixels is a function of the physical size of the marker and its distance from the camera. As shown in Table 3.1, the larger the marker size the further away the marker can be detected. A 50 mm marker – a size which could be displayed on a large number of mobile devices if using our phone-based marker technique – can be reliably captured at a distance of 1.4 m (sd = 0.09), which is a reasonable maximum distance to the center of the circle for conversational arrangement such as those in [46].

Table 3.1: Maximum ranges for different sized markers.

Marker Size (mm)	Maximum Range (mm)
50	1400
80	2300
100	2600
120	3200
150	3900

Marker tracking is also affected by the marker orientation relative to the camera. Markers were mounted on a table. The camera was moved at a constant distance of 1 m from the marker at varying rotation until the marker was no longer recognized. The maximum detectable angle with respect to vertical is approximately 75°, or almost vertical, meaning the camera should be able to resolve the angle from the marker in most comfortable-to-hold positions. Because of the highly accurate pose estimates, over an acceptable span and distance, the marker-based system is sufficient to provide automatic radial ordering.

3.6.2 Compass-Based Ordering

The abstract orientation sensor has a span over the entire 360° of the circle but is often characterized by noisy measurements, impacting the accuracy and precision of the position estimate. We recorded the reported orientation of eight phones arranged in a circular pattern as shown in Figure 3.14 for 15 minutes.

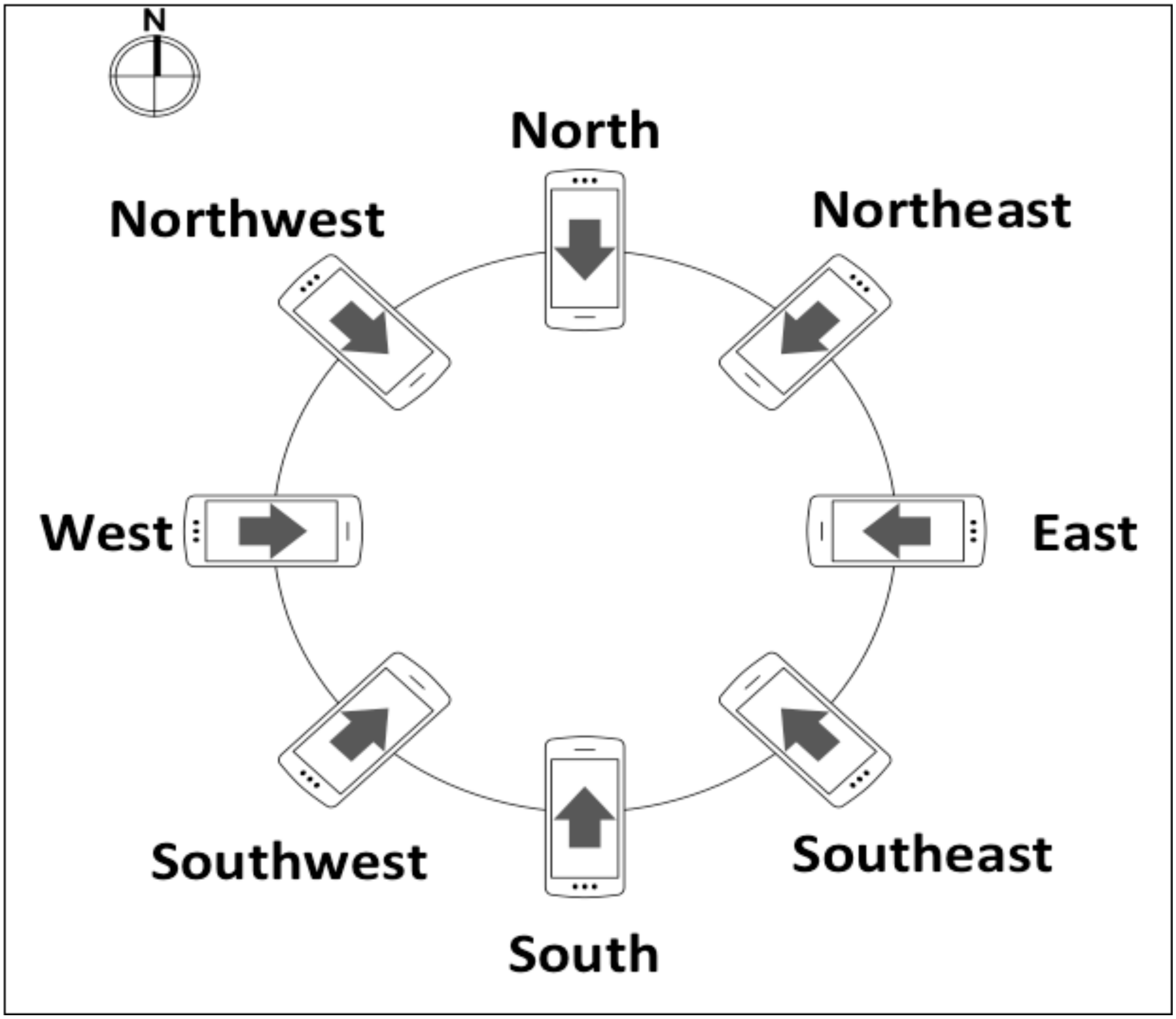


Figure 3.14: Experimental setup for compass-based ordering

Approximately 9000 data points were recorded on each phone. The test result shows that the distribution of the angle for each direction follows the Gaussian distribution, with a mean of -3°, which establishes more than sufficient accuracy. Angular error for each phone was calculated to determine the precision. Figure 3.15 shows the histogram of measurement of angle errors of all measurement angles. The 99% confidence interval lies at ±6.67°, allowing up to 54 people to be placed around the circle in the limit. Practically, a much smaller number will need to be localized. Based on our criteria, the compass has sufficient accuracy and precision, and span to provide the quality of service required.

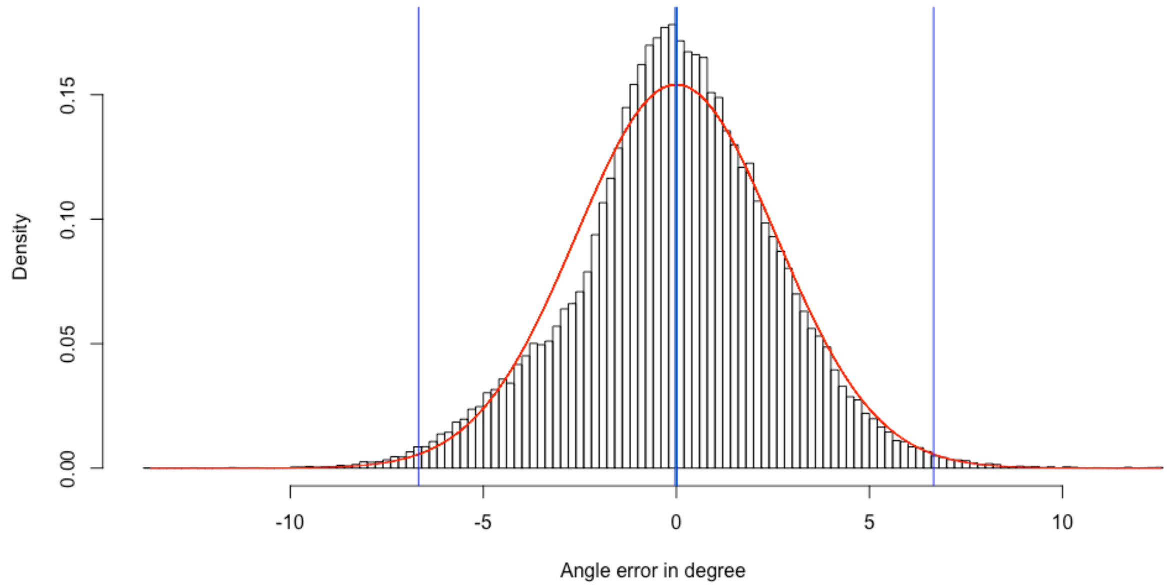


Figure 3.15: Histogram of measurement of angle errors (compass), red curve represents the best fit Gaussian, blue bars represent the 99% confidence interval.

3.6.3 Light-Gradient-Based Ordering

The light-gradient-based system works similar to our compass-based system. The difference is we use light direction as our "North". The same as the compass-based ordering technique, we recorded the reported orientation of eight phones arranged in a circular pattern for 15 minutes, collecting 33532 data points. We put a lamp bulb close to the surface of the table as a spot light source. Like the compass-based technique, the distribution of the angle for each direction follows a Gaussian distribution. Figure 3.16 shows the histogram of measurement of angle errors of all measurement angles. The 99% confidence interval lies at ±8.51°.

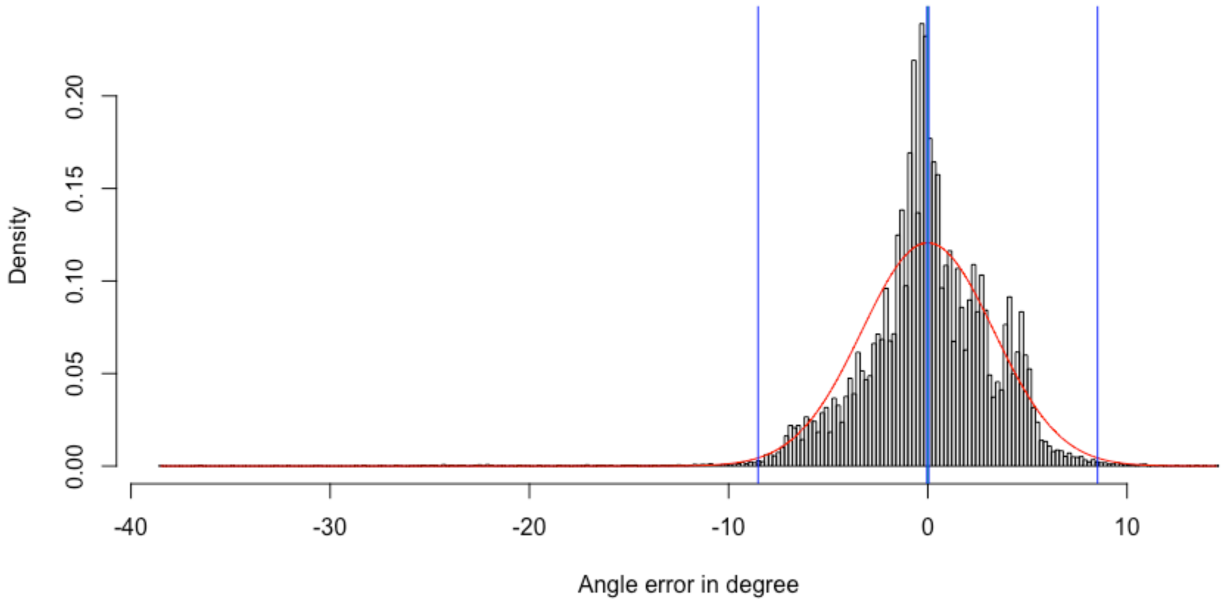


Figure 3.16: Histogram of measurement of angle errors (light-gradient), red curve represents the best fit Gaussian, blue bars represent the 99% confidence interval.

3.7 EXTENSIONS

3.7.1 File Transfer Application

We developed a file transfer Android application based on our compass-based auto-ordering technique (application UI shows in Figure 3.17a). A planet (Earth) represents the file to be sent. Remote devices are displayed as different planets. To send a file to a remote device, user simply put the "Earth" on the target planets. A connection will be established between the user's device and remote device. The file that the Earth represents will be sent to remote device via local Wi-Fi network.

The application's workflow is shown in Figure 3.17b.

(1) A file sending request is sent to the Bluetooth server which handles the device ordering information.

(2) The Bluetooth server sends a notification to the receiver device.

(3) Once the receiving user has accepted the file, the receiver device opens a server socket and notifies the Bluetooth server.

(4) The Bluetooth server broadcasts a server start notification to the sender device.

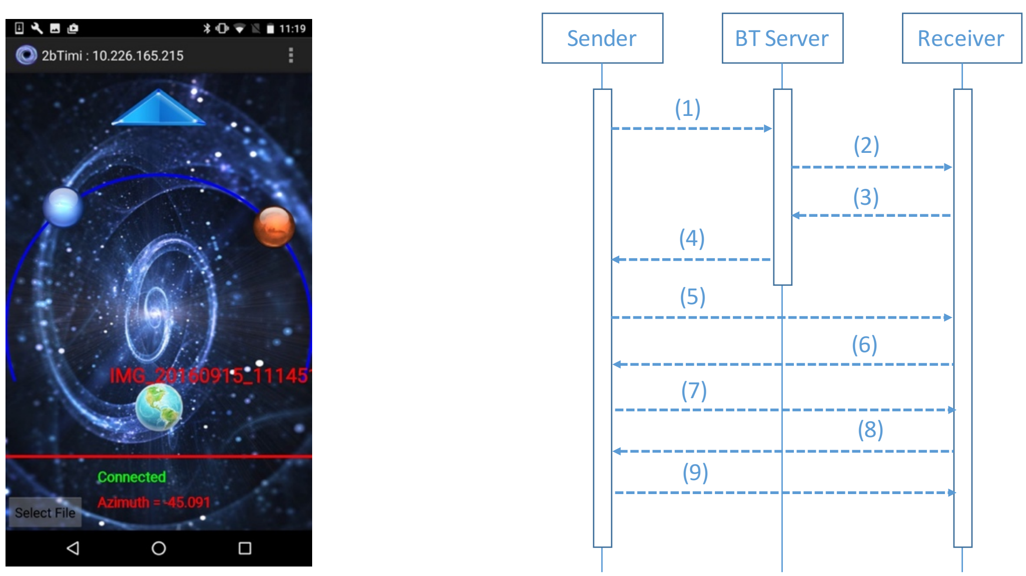
(5) When the sender device receives the server started notification, it will start to connect to the receiver device using the broadcast IP address and port information.

(6) When the connection is established, both the sender and the receiver devices get notifications from the Android framework.

(7) The sender device sends the file header information which contains the file name and file size to the receiver device.

(8) Once the receiver device receives the file header, it shows a progress dialog and notifies the sender device that the receiver is ready to accept the file.

(9) After sending the file header information and receiving confirmation from the receiver device, the sender device shows a file sending progress dialog and send the file. Once the file transfer is completed, the receiver device closes the server socket and the sender device disconnects from the receiver device.



(a) (b)

Figure 3.17: (a) application UI; (b) application workflow

3.7.2 Lock Mechanism

The lock mechanism enables user to stop updating position information. Basically we add a "I'm done localizing" status which “locks in” the location once the user is satisfied with their relative placement. Then the phone can be used for aiming, but only changing the relative location of the user, and not the targets. When the user performs an aiming task with the phone, it is mathematically equivalent to walking around the outside of the virtual circle to a point where users are opposite the target they are aiming at. If we don't lock the positions, then all users will see this virtual movement around the circle, which would be problematic if multiple users are moving simultaneously. However, if we lock positions, then only the local representation would change, allowing people to aim freely and simultaneously.

(a)

(b)

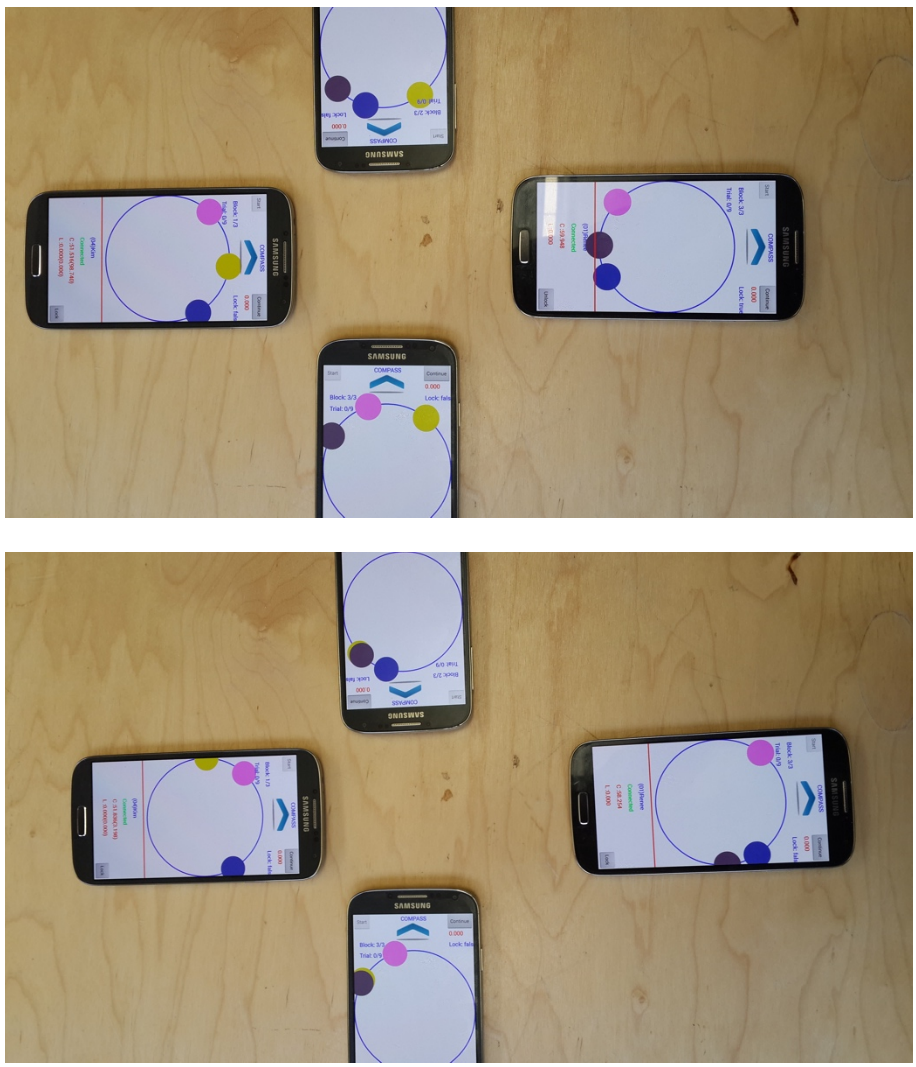


Figure 3.18: (a) static lock; (b) no lock

Two lock modes were developed. The first mode is Dynamic Lock. In this mode, location information is locked when user touch the file to be sent (e.g., the "Earth" in our file transfer application). The second mode is Static Lock. A lock button is provided on the screen. Location information is locked in when the button is pressed. Figure 3.18a shows the effect of our lock mechanism (Static Lock). The right phone is not pointing to the center of the circle. The location of remote devices still displayed correct under our static lock mode. In the unlocked mode (Figure 3.18b) the rotation of the right most phone was disrupted the representation on all others.

3.7.3 Bayesian Fusion

We developed a special auto-ordering technique which combines the compass-based ordering with the light-gradient-based ordering techniques. Both the compass sensor's reading angle and light gradient's reading angle can be fused in this technique. From previous technique assessment, the precision of compass and light gradient techniques are known, and follow a Normal distribution. The probability of each possible angle based on the current reading from each sensor and the error distribution (Equation (3.1)). The joint probability of compass and light gradient can be calculated by Equation (3.2). The angle with the maximum joint probability is a reasonable estimate of the final angle. Since the light-gradient-based technique is relatively unstable, the reading angle of light gradient will be used only if the light gradient is accurate, by setting a filter on light gradient data. When the standard deviation of the most recent 15 readings of light gradient is greater than 5 degrees, we consider the light gradient data to be inaccurate, and our fusion algorithm will only use the compass reading.

$$P(Y = i|X) = \frac{P(X|Y=i)P(Y=i)}{\sum_j P(X|Y=j)P(Y=j)} \tag{3.1}$$

$$P(X|Y = i) = P(X_{compass}|Y = i)P(X_{light-gradient}|Y = i) \tag{3.2}$$

CHAPTER 4

COMPARATIVE EVALUATION[1]

We carried out a controlled experiment to compare the performance of two of the auto-ordering techniques (Marker and Compass) to two existing approaches – unordered portals (which provide an on-screen proxy for each person, but ordered arbitrarily), and a proximity technique which detects when two devices are physically close (using phones' NFC radios). We did not test the gradient technique, as the current implementation of this method only works well in controlled laboratory conditions. A sufficiently robust image-based technique would appear identical to the marker-based technique, but without requiring the marker, which although advantageous would not impact the results of a usability study of the impact of radial auto ordering.

In all cases, a single phone was configured as the server, which maintained the known configuration of the phones and communications. Phones were dynamically added to the sharing representation as they logged into the app, and were connected to a Bluetooth star network via the server.

4.1 EXPERIMENTAL CONDITIONS

The transfer techniques were implemented in a simple experimental system that asked participants to transfer objects to one of three other people. Four techniques were developed for our experiment. They were Wormhole, Compass, Marker and Tap.

4.1.1 Wormhole

As seen in previous literature, our portal-based technique provided an on-screen proxy for each person; transfers were accomplished by dragging the object to the correct person's portal (Figure 4.1). Unlike the auto-ordered techniques, portals were ordered based on the time when devices connected to the server.

[1] Material in this chapter originally appeared in the following publication:
Li, C., Gutwin, C., Stanley, K., Nacenta, M.A. (2016). All Across the Circle: Using Auto-Ordering to Improve Object Transfer between Mobile Devices. In *Graphics Interface* (GI'2016), Victoria, BC, Canada. 49-56.

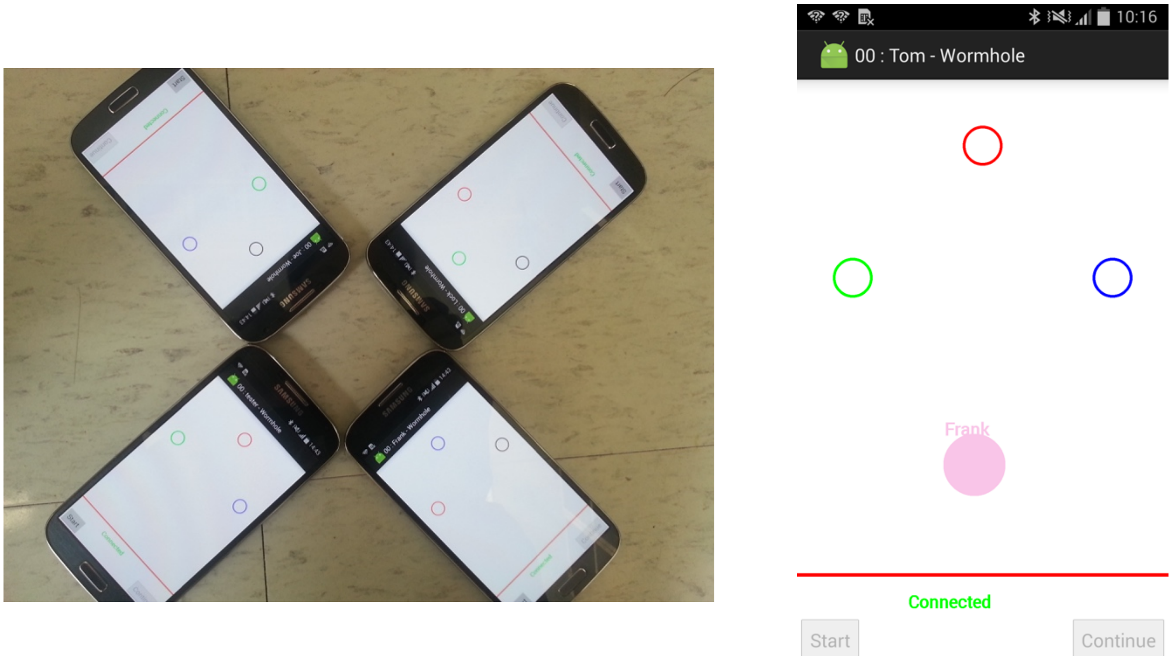


Figure 4.1: Wormhole

4.1.2 Compass

As described above, the Compass technique uses each device's compass reading to create a circular ordering for all devices (Figure 4.2). In order to make interactions as similar as possible across different conditions, the Compass technique used a proxy-based transfer action in which participants dragged objects to on-screen portals – although the portals were now ordered to match the locations of the other people around the circle. Our implementation can also use flick-based transfer, but only proxy-based transfer was used for the study, for internal consistency.

4.1.3 Marker

As described above, the Marker technique uses the device's camera to locate the fiducial marker and determine the orientation of the device to the marker; an ordering is then created using these relative orientations (Figure 4.3). As with Compass, the Marker technique used on-screen portals for transferring objects; the position of the portals was determined by the Marker-based ordering algorithm.

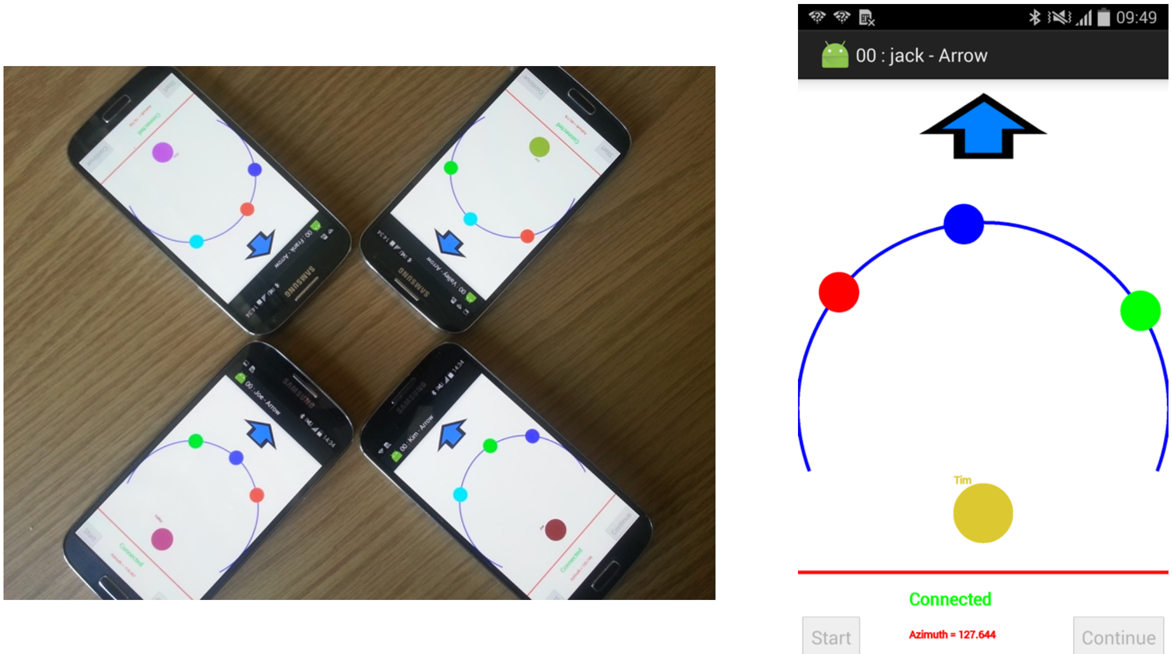


Figure 4.2: Compass

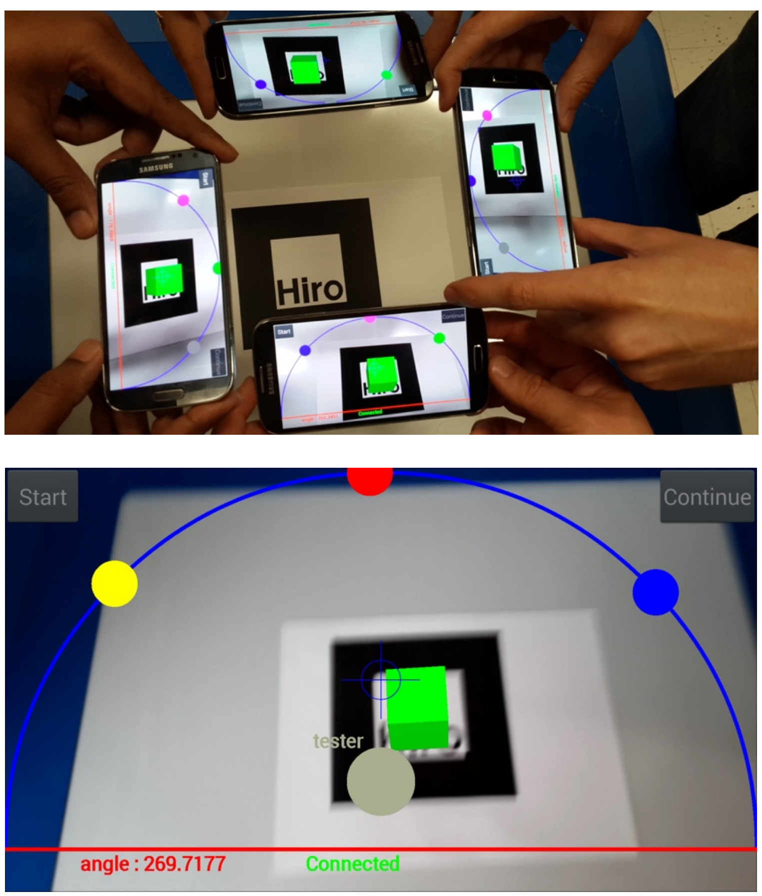


Figure 4.3: Marker

4.1.4 Tap

We developed a proximity-based transfer technique based on Hinckley's previous Bump system [30]. The Tap technique uses the Standard Android support for Near-Field Communication (NFC) to control the object transfer. To use this technique, participants held their devices back to back; when the devices were close enough, the sender saw a popup message on their display, and tapped the screen to complete the transfer over Bluetooth (see Figure 4.4).

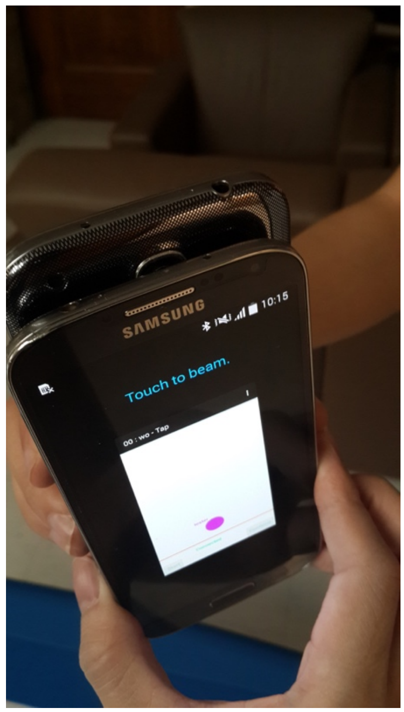


Figure 4.4: Tap

4.2 STUDY DESCRITION

4.2.1 Apparatus

The study used custom software developed with Android version 4.4.2, and was deployed on four Samsung Galaxy S4 devices (1.6 GHz processor, 5-inch 1080p display). All of the four experimental conditions were developed in the JAVA language, using Android Studio as the IDE. The study was conducted in an open area of a research lab (approximately 8m by 8m). The floor was marked with a two-meter-diameter circle and four locations at the north, east, south, and west points of the circle.

4.2.2 Participants

Thirty-two participants (12 female, 20 male) were recruited from a local university (ages 19-39, mean 26.6), in groups of four. Two of them were left-handed. All participants were frequent users of mobile devices (mean 23.6 hours/week). In seven of the eight groups, participants did not know one another.

4.2.3 Task

Participants were asked to transfer several objects to others in the group, using the mobile devices provided in this study. We simulated a setting where people transferred objects to an ad-hoc group that would be together for only a short time – such as an impromptu meeting at a conference coffee break. In order to simulate this setting (in which people know that they want to transfer to a particular person standing in the circle), we had participants wear nametags with made-up names, and for each transfer, a circle (representing the object to be transferred) appeared on the participant's device with the name of one of the other people in the group.

In the tap condition, they had to bring the phones in close proximity. For the Compass, Marker, and Wormhole techniques, the participant completed the transfer by dragging the circle to the correct on-screen portal. In all portal conditions, the portals were labeled with the participant's username (not their real name). We did not show usernames on the portals because this would have allowed simple pattern-matching between the named transfer object and because, in a real-world situation, the transfer object would not show the intended recipient. To allow people to build a memory mapping between portals and people in the Wormhole condition, names could be shown by long-pressing (500ms) anywhere on the screen.

4.2.4 Procedure

Each group of four participants completed demographic questionnaires, and then were given an introduction to the four transfer techniques after completing informed consent consistent with our ethics approval. Groups worked with each of the four transfer techniques in an order balanced using a Latin square. For each technique, participants carried out five blocks of trials. In each block, participants completed three transfers (trials) to each of the other people in the group (in random order) for a total of nine transfers.

To test the ability of participants to remap digital to physical locations when configurations change. After each block, the study simulated a new meeting of the four people – participants were moved to different physical locations around the circle, and the on-screen locations of the portals for all techniques except tap were reordered.

4.2.5 Measures

After each condition, participants completed an effort questionnaire based on the NASA Task Load Index (TLX) [26], asking about the technique they had just completed. After all four conditions, participants answered questions about their preferences.

The study used a 4×5×3 within-participant RM-ANOVA with factors *Transfer Technique* (Compass, Marker, Wormhole, Tap), *Block* (1-5), and *Repetition* to the same target (1,2,3 for each recipient). Dependent measures were transfer time (the time from the object appeared to the participant successful completed the object transfer task) and number of errors (the number of they send the object to the wrong target).

4.2.6 Hypotheses

We had the following hypotheses:

H1. Object transfer times for the auto-ordering techniques (*Compass* and *Marker*) will be faster than for either *Tap* or *Wormhole*;

H2. Error rate for *Compass* and *Marker* will be less than for *Wormhole*;

H3. Users will prefer the auto-ordering techniques over the other techniques.

4.3 RESULTS

4.3.1 Transfer Time

As shown in Figure 4.5, mean transfer times ranged from about two seconds for the auto-ordering techniques to above twelve seconds for Tap. Figure 4.6 displays only the three faster techniques, and shows that Wormhole was slower in the first two blocks, and then the same speed as the auto-ordering techniques.

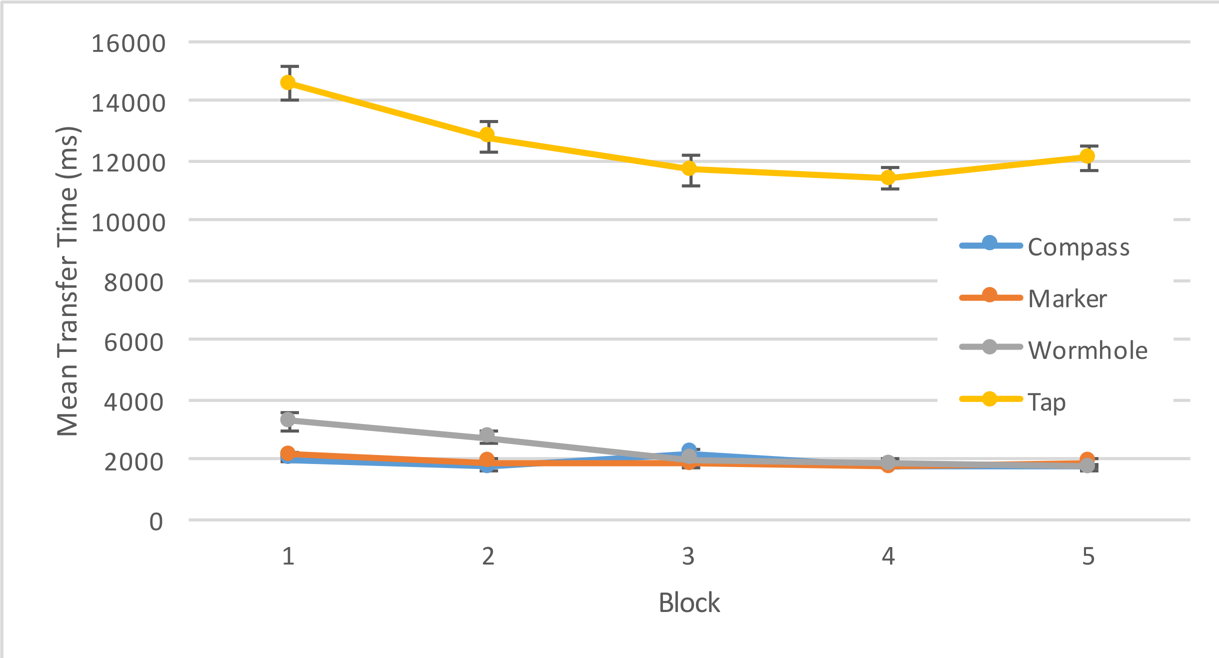


Figure 4.5: Mean transfer time (± s.e.), by technique and block

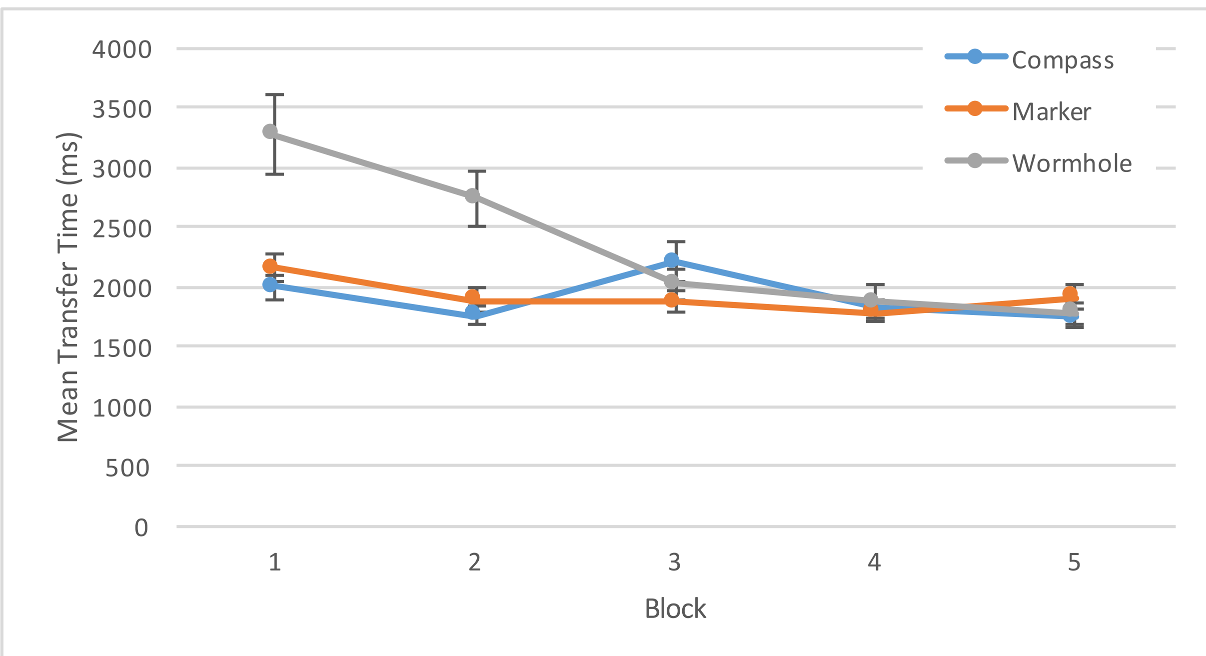


Figure 4.6: Mean transfer time (± s.e.) for portal techniques

RM-ANOVA showed a significant main effect of *Technique* ($F_{3,93}$=973.6, p<.0001), and also a significant interaction between *Technique* and *Block* ($F_{12,372}$=7.35, p<.0001). Follow-up pairwise comparisons between conditions (using Bonferroni correction to maintain alpha of 0.05) showed that *Tap* was slower than all other conditions, and that both *Compass* and *Marker* were faster than *Wormhole*. We therefore accept hypothesis H1.

Observations during the trials suggested that the main reason that *Tap* was slower than the other techniques is that people needed to wait for the other person to be ready to carry out the technique – that is, *Tap* requires both sender and receiver to engage in the transfer, whereas the other techniques allow unilateral transfer.

Observations and participant comments also suggested that the reason for slower performance of the *Wormhole* technique was that participants needed to remember the mapping between the portals and the people in the real world. Our analysis of the number of times targets were repeated provides additional insight into this issue. Figure 4.7 shows the transfer times for the three portal-based techniques on each of the three repetitions per block. *Wormhole* was slower on the first trial (when the locations were unknown), and then similar in speed for the second and third trials. RM-ANOVA showed a significant interaction between Technique and Repetition ($F_{6,186}$=11.43, p<.0001).

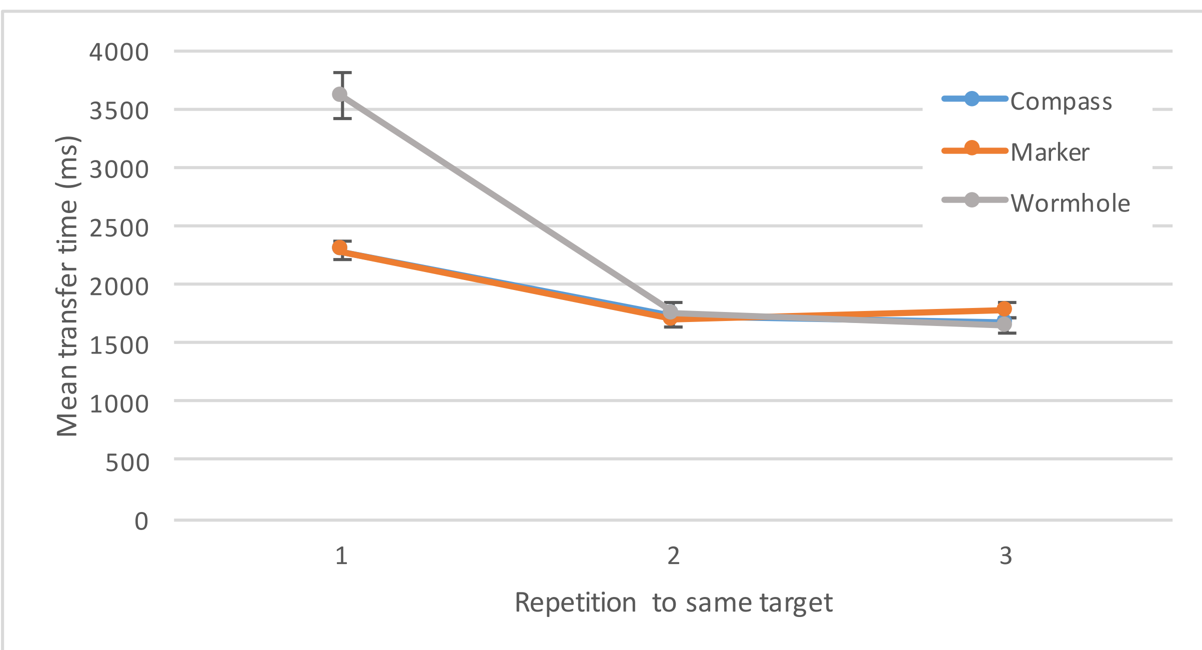


Figure 4.7: Mean transfer time by number of repetitions

4.3.2 Errors

We compared the number of errors made per transfer (an error was counted if the participant released the object on the wrong portal, or tapped devices with the wrong person). Figure 4.8 shows that the *Wormhole* technique had higher errors, particularly in the first two blocks. RM-ANOVA showed a significant main effect of Technique on errors ($F_{3,93}$=15.4, p<.0001), and also a significant interaction between Technique and Block ($F_{12,372}$=5.95, p<.0001). Follow-up pairwise comparisons show Wormhole had higher error rates than other techniques. We therefore accept hypothesis H2.

As shown in Figure 4.9, the number of errors for the *Wormhole* technique was highest on the first repetition (22%), falling to about 10% for the second and third trials. RM-ANOVA showed a significant interaction between Technique and Repetition ($F_{6,186}$=4.05, p<.005).

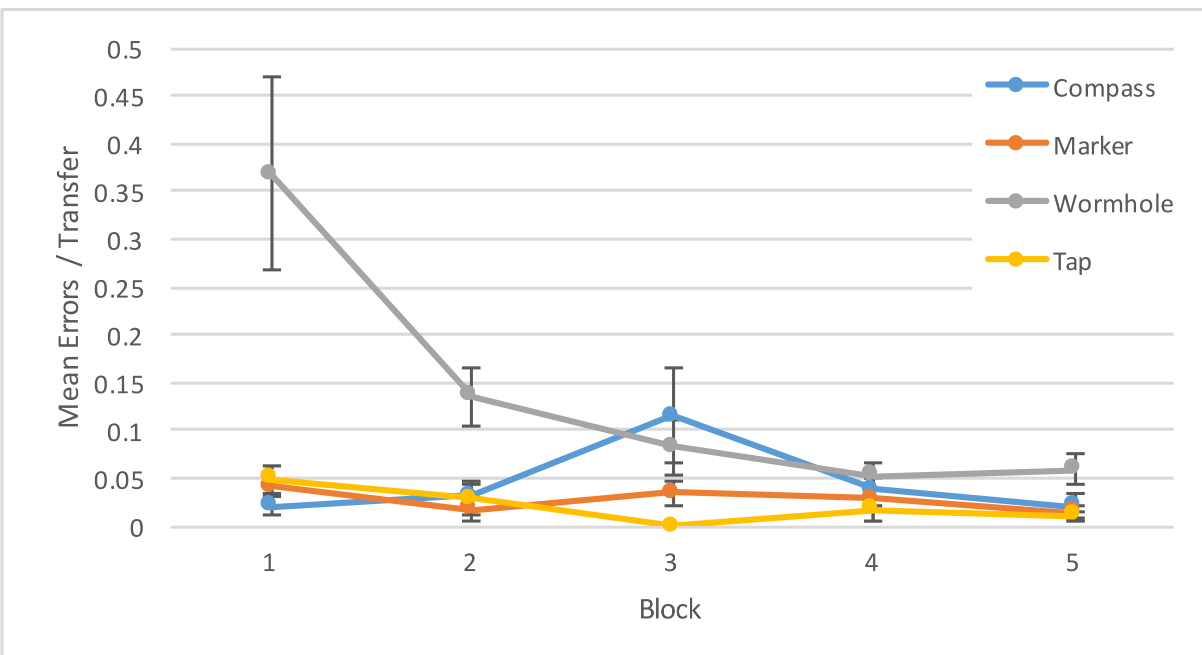


Figure 4.8: Error rate by technique and block

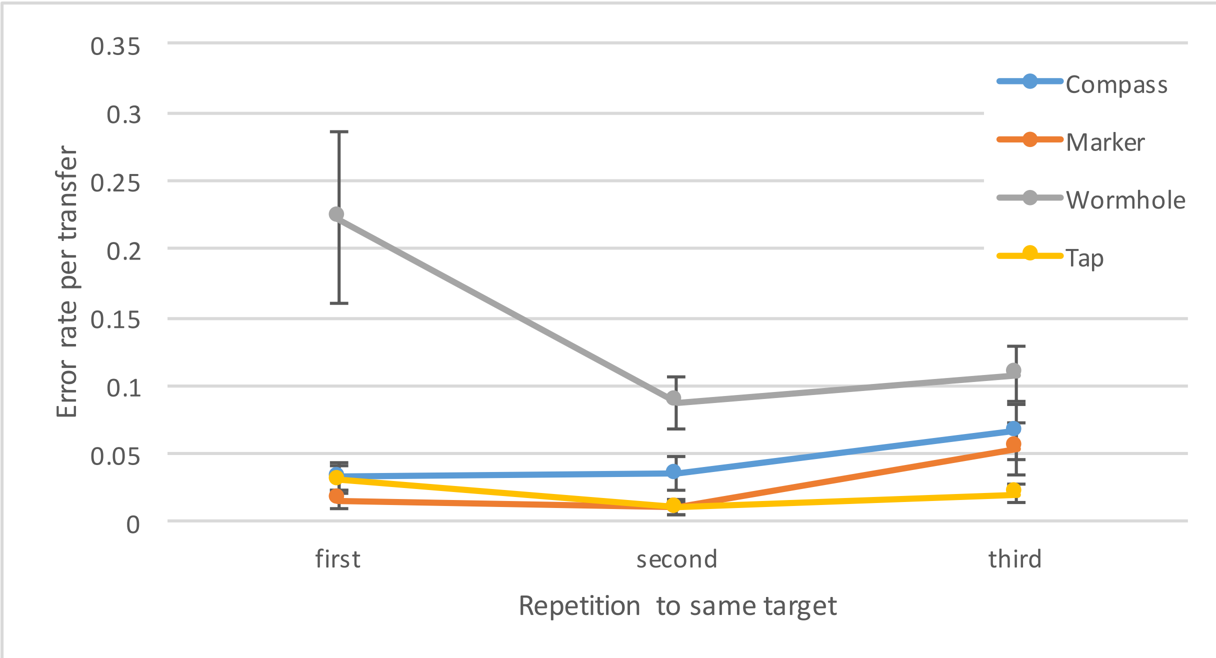


Figure 4.9: Error rate by repetition to same target

4.3.3 Subjective Response: Effort and Preferences

Responses to the post-condition questionnaire (based on the NASA-TLX) are shown in Table 4.1. Friedman tests on the ratings showed that there were significant differences between the techniques in the amount of mental and physical effort, the amount of work required, and the level of frustration (all $p<0.05$). For all questions, ratings of the *Tap* condition were highest; the other three conditions (*Wormhole*, *Marker*, and *Compass*) were similar.

We also asked participants which technique they preferred in terms of several qualities – ease of use, speed, accuracy, and overall preference (see Table 4.2). Chi-squared tests showed that significantly more participants chose *Compass* for ease of use, speed, and overall preference. Most participants also chose *Compass* for accuracy, but the test was not significant. We therefore accept hypothesis H3.

Table 4.1: Mean (s.d.) of effort scores (1-7 scale, low to high)

	Compass	Marker	Wormhole	Tap	χ^2	*p*
Mental	2.2 (1.4)	2.6 (1.4)	3.3 (1.7)	3.1 (1.7)	9.75	**.021**
Physical	2.1 (1.5)	2.8 (1.8)	2.0 (1.2)	4.3 (1.9)	23.3	**<.0001**
Temporal	3.4 (1.9)	3.1 (1.6)	3.4 (1.8)	3.9 (1.6)	4.43	0.21
Success	6.2 (1.1)	6.0 (1.4)	5.7 (1.2)	5.4 (1.8)	5.87	0.12
Hard work	2.6 (1.9)	2.8 (1.8)	3.4 (1.8)	4.1 (1.9)	10.9	**.012**
Frustration	2.0 (1.4)	2.6 (1.9)	2.3 (1.5)	3.6 (2.1)	12.2	**.0068**

Table 4.2: Counts of participant preferences

Which was:	Compass	Marker	Wormhole	Tap	χ^2	*p*
Easiest to use	21	5	3	3	28.5	**.0001**
Fastest	19	7	5	2	20.2	**.0002**
Most accurate	14	6	5	8	5.91	0.12
Overall preference	16	7	6	3	11.7	**.008**

4.3.4 Participant Comments

We asked participants to provide written comments after each condition, and their remarks follow the performance and preference results provided above.

First, several people commented on the speed of *Tap*, and particularly the need to wait for the other person to be ready. For example, one participant said “We have to wait until the person that I had to send the ball to is available to tap. Therefore, it is not that independent;” another stated “Very annoyed with having to wait for other people to transfer.” On the other hand, although Tap was time consuming participant considered Tap was a technique that is good for secure transmission. A participant said “it is secured as it involves two people.” A few people stated that physical proximity with other person is an interesting technique. For instance, a participant commented “This one was most interesting, as you had to work with others to softly agree whose turn it was, and in what order you'd go.” Another participant said “I think it could make the participant more interested in get touch with others.”

Second, participants recognized the correspondence between people in the real world and on the screen (for the two auto-ordering techniques) and the lack of correspondence for the *Wormhole* technique. Regarding *Wormhole*, one person said “I had to touch the circle to see who they were first before sending the ball.” Another thought “It was little bit tedious to remember the location.” For *Marker*, one participant said “The exact positions were represented on the mobile device, hence it was easier to locate;” another stated that *Compass* was easy “because I didn't have to remember the positions;” and a third said (also about Compass) “It was just the most straightforward of them all, and I didn't have to remember any particular order to the positioning, as I could rely on both the screen or the person's name badge 'in real life'.”

Third, for the two auto-ordering techniques. Participants considered the Marker technique is more cumbersome than the Compass technique. One participant said “It is similar to arrow, however keeping the cube in the center is kind of irritating;” and another commented “nice thing about this task was the circles were stable but I had problem keeping camera stable and also moving the object to the ball.”

We ask participants to state their overall preference after all conditions. Half of the participant chose *Compass* as their most preference technique. They considered “it was the simplest to interpret and use, and the fastest." Another participant said "It was easy to recognize the people around in this condition. More simple instructions."

From what has been discussed above, we consider our auto-ordering techniques, especial *Compass* technique, are superior other techniques on user performance and preference when doing object transfer tasks. We accept all of our three hypotheses. Our auto-ordering techniques are faster than two traditional techniques; our auto-ordering techniques have less errors than standard portal-based technique; most of the participants prefer to use our auto-ordering techniques.

CHAPTER 5

DISSCUSSION, CONTRIBUTIONS AND CONCLUSTIONS

This chapter provides a discussion of the findings of our comparative experiment. This chapter also presents several limitations to our auto-ordering techniques and recommendations for future work. We also introduce the contributions that are made by this thesis, and finish with a final summary.

5.1 DISCUSSION

We developed and tested three new techniques for auto-ordering devices that are in an approximate circle (an F-Formation). We demonstrated the technical capabilities of the three techniques, and carried out a comparative study using two of our techniques (one based on the compass, and one based on a fiducial marker). The main findings of the study were:

- Auto-ordering techniques were faster than *Tap*;
- Auto-ordering techniques had lower error rates than *Wormhole*, particularly at first;
- *Compass* was strongly preferred by users.

The findings from our evaluation generally match our expectations of the capabilities and limitations of the techniques and their underlying sensing technologies.

First, the slow speed of the *Tap* technique appears to be caused by the requirement that both participants (sender and receiver) participate in the transfer gesture. Because participants were often engaged in a transfer to another person, this requirement meant that people spent considerable time waiting for the receiver to be available to match the tap gesture. This disadvantage makes *Tap* does not apply to the situation like business cards sharing among a group of people during coffee break of a conference. Although this delay would not always occur (e.g., in a single-transfer scenario), there is a performance advantage for “one-sided” techniques that allow a transfer to be carried out with only one person.

The performance of auto-ordering compared to *Wormhole* appears to arise from these techniques’ correspondence between the real world and the on-screen representation of targets. As noted several times by our participants, it was easier to carry out the transfer when the action was guided by the real world as well as the on-screen target. The higher error rate of *Wormhole* has a similar explanation – the arrangement of targets had to be memorized in each block, and when the arrangement did not match the real world, people had to deal with conflicting information.

We note that the correspondence problem is reduced for the *Wormhole* technique in some situations. For example, when people already know the names or IDs of the people around the circle, then the labelling of portals will provide enough information for people to carry out the transfer, without needing to build any memory mapping, for example a family group, or a workgroup that knows each other well. Second, *Wormhole* can work as well as the auto-ordering techniques when positions remain stable. As shown in Figure 4.7, people's performance with *Wormhole* got closer to *Compass* in the second and third repetitions to each target, likely because people were able to memorize the mapping. In addition, Figure 4.6 shows that people were also faster with this technique in the later blocks of the study, suggesting that people learned how to best use the technique. However, research on Stimulus-Response compatibility has shown that there are performance advantages in using spatially compatible arrangements even after extensive training [56].

However, even though *Wormhole* can work well in some situations, our study shows that this technique has poor performance when people are dependent on the real-world arrangement of the group – e.g., situations where people do not know one another. Our auto-ordering techniques do not suffer from this limitation, and perform no worse when people are familiar.

Finally, the preference for the *Compass* over *Marker* is likely due to the reduced constraints on how participants had to hold their devices – with *Marker*, people had to keep the fiducial marker in the camera's view while they carried out the transfers, whereas with *Compass* they had much more freedom to hold the device as they wished, as long as it generally pointed towards the center of the circle. Participants' comments also confirmed this. They thought it is irritating to keep tracking the fiducial marker. They considered the *Compass* technique was more intuitive and easier to use than the *Marker* technique.

5.2 CONTRIBUTIONS

This thesis makes three primary contributions: we have developed new localization techniques for effective object sharing, but in doing so have also added to the literature by providing comparative analysis and a further exploration of F-Formations.

- *New Localization Techniques*: Our primary contribution in this work is three new localization techniques for determining device ordering. Our techniques allow fast and accurate object transfer compared to two standard approaches (unordered portal-based and proximity-based techniques). Our auto-ordering techniques are based on sensors and computational resources readily available on almost all mobile devices, showing that our techniques are almost immediately usable. The technical evaluation demonstrated that the sensor accuracy, precision and span were more than sufficient for typical use. Our techniques were overwhelmingly preferred by participants, indicating a strong potential for uptake. The Compass technique in particular is immediately usable on most smartphones, and should allow future research in ad-hoc message or file passing to be conducted simply, cheaply and reliably.
- *Comparative Evaluation*: Although our techniques were superior in many ways, they may not always be appropriate or possible. Designers now have empirical evidence concerning the tradeoffs between the different techniques. Tap is slow, but had close to zero errors, and could be useful when security, and in particular, recipient selection is of paramount importance. For instance, Tap would be a good fit for a money transfer task. Our analysis was the first to demonstrate Tap's inherent timing disadvantage anchored in the requirement for mutual action. Wormhole always performed the worst initially, but had its performance converge to that of our techniques within two repetitions. For tasks where repeated transfers have to take place wormhole might be superior as it can be performed without sensing, such as an office environment where all devices' locations are fixed. Furthermore, we were the first to demonstrate that in most of the cases discussed in the research of Marquardt et al. [46], our techniques would be preferred.
- *Theoretical Grounding*: Our design draws heavily on the idea of F-Formations, which shows the spatial arrangements that people typically adopt in ad-hoc groups. The success of our techniques from both a performance and preference perspective provides additional

support to the validity and utility of F-Formations as a construct for designing co-located collaborative systems. Furthermore, pointing to the circle is necessary for calibration but not necessary if people's locations are reasonably stable. For example, a configuration screen could be used as a setup step. Once the users are satisfied with their relative placement, our lock mechanism, introduced in Chapter three, could enable people to aim freely and simultaneously.

Beyond our immediate contributions, our work could have significant impact on other areas. The most obvious and immediate application of our work is as an interface widget in other file sharing studies. Given the apparent superiority of the technique, individuals studying other aspects of file sharing amongst collocated handheld devices (for example preview modes for received files) should adopt our technique for ordering to minimize the timing and learning confounds found in tap or wormhole techniques, respectively.

While initially designed to solve the problem of auto-ordering for file sharing in ad-hoc collaborative groups, the technology has the potential for integration into larger collaborative systems; for example, facilitating file sharing amongst cliques of groups in a conference or work environment.

Finally, our technology serves as a demonstration of how simple spatial sensors, now ubiquitously available on smartphones can be used in clever ways to facilitate collaborative actions. These kinds of interaction techniques might be interesting in co-located games, for example passing a virtual hot potato or as part of a live action game of "Simon." Further research on other techniques in the area could revolutionize environments as diverse as the office, and virtual reality games.

5.3 LIMITATIONS AND FURTURE WORK

Although we have made novel contributions in this work, there remain some limitations to the study and a great deal of potential future work.

Our work is heavily dependent on the use of sensors standard on commodity smartphones. While these devices are generally reliable, the sensors do have well known failure modes. Smartphone compasses can provide noisy or unreliable readings in ferrous environments. This limitation does not overly constrain the number of possible use cases, however, and the sensor has been robust in our tests. Furthermore, this issue being actively addressed by sensor scientists, and may be overcome in 3-5 years. In future work we plan to empirically test the robustness of the techniques in real-world settings and with real-world groups. The evolution of cameras mounted on smartphones is rapid. We believe that the span of our marker-based auto-ordering technique could be improved. That is, a small marker could be recognized at further range, and a wider recognizable angle could be achieved.

Our usability evaluation should also be followed up by further studies in a more natural, less controlled environments, and should include a greater diversity of experimental tasks. It would also be desirable to test the limitations of the systems for number of simultaneous users, speed of transfer and stability of spatial arrangement. The work here is an important step in leveraging new technology for new collaborative techniques.

Finally, our implementations of portal-based techniques were intentionally limited (e.g., users could not rearrange portals to match the real world and the labels were invisible); in future work we will test whether these added capabilities could improve the overall performance of the Wormhole method.

5.4 CONCLUSIONS

By exploiting the common arrangements of individuals in small group gatherings, we were able to simplify a complex multi-agent spatial localization problem, to localizing agents around an approximate circle. We were able to leverage sensors available on commodity smartphones to localize individuals around a circle with a resolution sufficient for at least twenty-two people, which is more than what is currently supported for Bluetooth ad-hoc networks. We developed three techniques: a marker-based technique which can use either a paper or phone-based marker, a compass-based interaction which had lower sensed precision but works without a calibration step, and a vision-based technique that senses the gradient of the background. In a controlled experiment, auto-ordering techniques consistently outperformed unordered portals and a proximity-based technique, and were preferred by most participants. Because these techniques employ standard smartphones, they can be easily deployed, and can help facilitate digital object sharing in small group environments.

5.4.1 Summary of Thesis

This thesis focused on the development of three new localization techniques for determining device ordering. The first of our auto-ordering techniques uses the device's camera to locate a fiducial marker and determine the orientation of the device to the marker and then uses this orientation to infer each phone's relative location around a circle. The second technique employs the device's compass to determine the relative angle with respect to magnetic North. An ordering is formed when users point their phones toward a location at the center of the circle. Our third technique uses device's camera to capture an image of the background. After applying image processing algorithms to the image, the brightness gradient of the background could be calculated. Devices' ordering around the circle can be resolved by the absolute direction given by the light gradient.

Three extensions were developed to provide additional functionality. The first is an Android file transfer application, which allows local file transfer. The second extension introduces a lock mechanism to our system. This mechanism relaxes the constraints of the F-formation system. A sensor fusion auto-ordering technique, which combines the compass-based and the light-gradient based techniques, is the third extension.

Two assessments were performed. The first was a technical assessment that measured the sensitivity, precision and span of the localization techniques. The second assessment was a usability study, which compared the performance of two of the auto-ordering techniques (Marker and Compass) to two existing approaches. Object transfer time and error rate were measured in this comparative study. The results of these two assessments indicated that our auto-ordering techniques were effective enough and usable to be used for object transfer tasks in small groups.

5.4.2 Concluding Remarks

The problem addressed in this thesis was that there is a lack of simple, intuitive and reliable localization techniques for determining device ordering. The main motivation for developing a localization technique for device ordering is to improve the efficiency of data-sharing. The solution presented here was to provide simple, intuitive and reliable localization techniques by using sensors standard on commodity mobile devices. We developed three systems that can auto-order locations about a circle based on different sensors commonly available on today's smartphones.

We conducted technical assessments to all of our systems and usability assessment to two of our techniques.

Our technical assessment showed that our solutions could provide sufficient accuracy and precision to reliably localize people around a circle.

The results of our usability evaluation found:

- Auto-ordering was faster than the other techniques, and less error prone than the portal technique.
- While participants were able to reduce transfer time using portals as they learned mappings, they never achieved better performance than the auto-ordering techniques.
- Participants overwhelmingly preferred the auto-ordering techniques to the portal and proximity techniques.
- The smartphone sensors underlying the techniques are accurate enough for groups of up to twenty people – many more than will typically be encountered in ad-hoc groups.

Our techniques provide a simple, intuitive and reliable solution for a common transfer situation — a small group gathered in an approximate circle — with sufficient accuracy and precision to reliably localize people around a circle.

REFERENCES

[1] Android Development Documentation. URL: https://developer.android.com/index.html

[2] ARToolKit. URL: http://www.hitl.washington.edu/artoolkit/

[3] Ascension Technology Corporation. Website: http://www.ascension-tech.com/products/#product-history

[4] Azuma, R. T. (1997). A survey of augmented reality. Presence: Teleoperators and virtual environments, 6(4), 355-385.

[5] Bahl, P., & Padmanabhan, V. N. (2000). RADAR: An in-building RF-based user location and tracking system. In *INFOCOM 2000. Nineteenth Annual Joint Conference of the IEEE Computer and Communications Societies. Proceedings. IEEE* (Vol. 2, pp. 775-784). Ieee.

[6] Banerjee, N., Agarwal, S., Bahl, P., Chandra, R., Wolman, A., & Corner, M. (2010, May). Virtual compass: relative positioning to sense mobile social interactions. In *International Conference on Pervasive Computing* (pp. 1-21). Springer Berlin Heidelberg.

[7] Baudisch, P., Cutrell, E., Hinckley, K., & Gruen, R. (2004, April). Mouse ether: accelerating the acquisition of targets across multi-monitor displays. In*CHI'04 extended abstracts on Human factors in computing systems* (pp. 1379-1382). ACM.

[8] Bell, S., Jung, W. R., & Krishnakumar, V. (2010, November). WiFi-based enhanced positioning systems: accuracy through mapping, calibration, and classification. In Proceedings of the 2nd ACM SIGSPATIAL International Workshop on Indoor Spatial Awareness (pp. 3-9). ACM.

[9] Benko, H., & Feiner, S. (2005, April). Multi-monitor mouse. In *CHI'05 extended abstracts on Human factors in computing systems* (pp. 1208-1211). ACM.

[10] Biehl, J. T., & Bailey, B. P. (2004, May). ARIS: an interface for application relocation in an interactive space. In *Proceedings of Graphics interface 2004*(pp. 107-116). Canadian Human-Computer Communications Society.

[11] Bluetooth SIG, the Bluetooth specifications https://www.bluetooth.com/specifications

[12] Bolt, R. A. (1980). *"Put-that-there": Voice and gesture at the graphics interface* (Vol. 14, No. 3, pp. 262-270). ACM.

[13] Booth, K. S., Fisher, B. D., Lin, C. J. R., & Argue, R. (2002, October). The mighty mouse multi-screen collaboration tool. In *Proceedings of the 15th annual ACM symposium on User interface software and technology* (pp. 209-212). ACM.

[14] Chen, X. A., Schwarz, J., Harrison, C., Mankoff, J., & Hudson, S. (2014, September).

Around-body interaction: sensing & interaction techniques for proprioception-enhanced input with mobile devices. In *Proceedings of the 16th international conference on Human-computer interaction with mobile devices & services* (pp. 287-290). ACM.

[15] Ciolek, T. M., & Kendon, A. (1980). Environment and the spatial arrangement of conversational encounters. *Sociological Inquiry*, *50*(3-4), 237-271.

[16] Cristani, M., Bazzani, L., Paggetti, G., Fossati, A., Tosato, D., Del Bue, A., ... & Murino, V. (2011). Social interaction discovery by statistical analysis of F-formations. In *BMVC* (Vol. 2, p. 4).

[17] Dearman, D., Guy, R., & Truong, K. (2012, May). Determining the orientation of proximate mobile devices using their back facing camera. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 2231-2234). ACM.

[18] Domhan, T., TIT07INA, K., & der Dualen Hochschule, G. (2010). Augmented reality on android smartphones. Studiengangs Informationstechni. Dualen Hochschule Baden-Württemberg Stuttgart.

[19] eDrawings. URL: http://www.edrawingsviewer.com/

[20] Edward, H. (1966). The hidden dimension. *Doubleday, Garden City*, *14*, 103-124.

[21] Gan, T., Wong, Y., Zhang, D., & Kankanhalli, M. S. (2013, October). Temporal encoded F-formation system for social interaction detection. In *Proceedings of the 21st ACM international conference on Multimedia* (pp. 937-946). ACM.

[22] Getting, I. A. (1993). Perspective/navigation-the global positioning system. *IEEE spectrum*, *30*(12), 36-38.

[23] Goel, M., Lee, B., Islam Aumi, M. T., Patel, S., Borriello, G., Hibino, S., & Begole, B. (2014, April). SurfaceLink: using inertial and acoustic sensing to enable multi-device interaction on a surface. In *Proceedings of the 32nd annual ACM conference on Human factors in computing systems* (pp. 1387-1396). ACM.

[24] Grossman, T., & Balakrishnan, R. (2005, April). The bubble cursor: enhancing target acquisition by dynamic resizing of the cursor's activation area. In *Proceedings of the SIGCHI conference on Human factors in computing systems* (pp. 281-290). ACM.

[25] Hamilton, P., & Wigdor, D. J. (2014, April). Conductor: enabling and understanding cross-device interaction. In *Proceedings of the 32nd annual ACM conference on Human factors in computing systems* (pp. 2773-2782). ACM.

[26] Hart, S. G., & Staveland, L. E. (1988). Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. *Advances in psychology*, *52*, 139-183.

[27] Harter, A., Hopper, Andy., Steggles, Pete., Ward, Any., and Webster, Paul. (1999) The anatomy of a context-aware application. *In Proceedings of the 5th Annual ACM/IEEE International Conference on Mobile Computing and Networking (Mobicom 1999* (pp. 59-68). ACM.

[28] Hascoët, M. (2008, April). Throwing models for large displays. In *HCI'03: 11th International Conference on Human Computer Interaction* (Vol. 2, pp. pp-73). British HCI Group.

[29] Hightower, J., & Borriello, G. (2001). A survey and taxonomy of location systems for ubiquitous computing. *IEEE computer*, *34*(8), 57-66.

[30] Hinckley, K. (2003). Bumping objects together as a semantically rich way of forming connections between ubiquitous devices. *Ubicomp2003*.

[31] Hinckley, K., Ramos, G., Guimbretiere, F., Baudisch, P., & Smith, M. (2004, May). Stitching: pen gestures that span multiple displays. In *Proceedings of the working conference on Advanced visual interfaces* (pp. 23-31). ACM.

[32] Holmquist, L. E., Mattern, F., Schiele, B., Alahuhta, P., Beigl, M., & Gellersen, H. W. Smart-Its friends: A technique for easily establish connections between smart artifacts. In *Proceedings of UBICOMP 2001*.

[33] Hung, H., & Kröse, B. (2011, November). Detecting f-formations as dominant sets. In *Proceedings of the 13th international conference on multimodal interfaces* (pp. 231-238). ACM.

[34] Jin, H., Holz, C., & Hornbæk, K. (2015, November). Tracko: Ad-hoc mobile 3d tracking using bluetooth low energy and inaudible signals for cross-device interaction. In *Proceedings of the 28th Annual ACM Symposium on User Interface Software & Technology* (pp. 147-156). ACM.

[35] Johanson, B., Ponnekanti, S., Sengupta, C., & Fox, A. (2001, September). Multibrowsing: Moving web content across multiple displays. In *International Conference on Ubiquitous Computing* (pp. 346-353). Springer Berlin Heidelberg.

[36] Kato, H., & Billinghurst, M. (1999). Marker tracking and hmd calibration for a video-based augmented reality conferencing system. In *Augmented Reality, 1999.(IWAR'99) Proceedings. 2nd IEEE and ACM International Workshop on*(pp. 85-94). IEEE.

[37] Kendon, A. (1976). The F-formation system: The spatial organization of social encounters. *Man-Environment Systems*, *6*, 291-296.

[38] Kendon, A. (1990). Conducting Interaction: Patterns of Behavior in Focused Encounters. Cambridge University. *Studies in Interactional Sociolinguistics*, (7).

[39] Kortuem, G., Kray, C., & Gellersen, H. (2005, October). Sensing and visualizing spatial relations of mobile devices. In *Proceedings of the 18th annual ACM symposium on User interface software and technology* (pp. 93-102). ACM.

[40] Kray, C., Nesbitt, D., Dawson, J., & Rohs, M. (2010, September). User-defined gestures for connecting mobile phones, public displays, and tabletops. In *Proceedings of the 12th international conference on Human computer interaction with mobile devices and services* (pp. 239-248). ACM.

[41] Kray, C., Rohs, M., Hook, J., & Kratz, S. (2009). Bridging the gap between the Kodak and the Flickr generations: A novel interaction technique for collocated photo sharing. *International Journal of Human-Computer Studies*,*67*(12), 1060-1072.

[42] Li, M., & Kobbelt, L. (2012, December). Dynamic tiling display: building an interactive display surface using multiple mobile devices. In *Proceedings of the 11th International Conference on Mobile and Ubiquitous Multimedia* (p. 24). ACM.

[43] Lloyd, D., Benford, S., & Greenhalgh, C. (1999, December). Formations: explicit group support in collaborative virtual environments. In *Proceedings of the ACM symposium on Virtual reality software and technology* (pp. 162-163). ACM.

[44] Lowe, D. G. (1999). Object recognition from local scale-invariant features. In *Computer vision, 1999. The proceedings of the seventh IEEE international conference on* (Vol. 2, pp. 1150-1157). Ieee.

[45] Marquardt, N., Diaz-Marino, R., Boring, S., & Greenberg, S. (2011, October). The proximity toolkit: prototyping proxemic interactions in ubiquitous computing ecologies. In *Proceedings of the 24th annual ACM symposium on User interface software and technology* (pp. 315-326). ACM.

[46] Marquardt, N., Hinckley, K., & Greenberg, S. (2012, October). Cross-device interaction via micro-mobility and f-formations. In *Proceedings of the 25th annual ACM symposium on User interface software and technology* (pp. 13-22). ACM.

[47] Marshall, P., Rogers, Y., & Pantidi, N. (2011, March). Using F-formations to analyse spatial patterns of interaction in physical environments. In *Proceedings of the ACM 2011 conference on Computer supported cooperative work* (pp. 445-454). ACM.

[48] Milgram, P., Takemura, H., Utsumi, A., & Kishino, F. (1995, December). Augmented reality: A class of displays on the reality-virtuality continuum. In Photonics for industrial applications (pp. 282-292). International Society for Optics and Photonics.

[49] Nacenta, M. (2009). *Cross-display object movement in multi-display environments* (Doctoral dissertation, Ph. D. Dissertation, University of Saskatchewan).

[50] Nacenta, M. A., Gutwin, C., Aliakseyeu, D., & Subramanian, S. (2009). There and back

again: Cross-display object movement in multi-display environments. *Human–Computer Interaction*, *24*(1-2), 170-229.

[51] Nacenta, M. A., Sallam, S., Champoux, B., Subramanian, S., & Gutwin, C. (2006, April). Perspective cursor: perspective-based interaction for multi-display environments. In *Proceedings of the SIGCHI conference on Human Factors in computing systems* (pp. 289-298). ACM.

[52] OpenCV library: http://opencv.org/about.html

[53] Ozuysal, M., Fua, P., & Lepetit, V. (2007, June). Fast keypoint recognition in ten lines of code. In *2007 IEEE Conference on Computer Vision and Pattern Recognition* (pp. 1-8). Ieee.

[54] Polhemus Corporation. Website: http://polhemus.com

[55] Priyantha, N. B., Chakraborty, A., & Balakrishnan, H. (2000, August). The cricket location-support system. In *Proceedings of the 6th annual international conference on Mobile computing and networking* (pp. 32-43). ACM.

[56] Proctor, R. W., and Reeve, T. G. 1989. *Stimulus-response compatibility: An integrated perspective*. Elsevier.

[57] Rädle, R., Jetter, H. C., Schreiner, M., Lu, Z., Reiterer, H., & Rogers, Y. (2015, April). Spatially-aware or spatially-agnostic?: Elicitation and evaluation of user-defined cross-device interactions. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems* (pp. 3913-3922). ACM.

[58] Rekimoto, J. (1997, October). Pick-and-drop: a direct manipulation technique for multiple computer environments. In *Proceedings of the 10th annual ACM symposium on User interface software and technology* (pp. 31-39). ACM.

[59] Schwarz, J., Klionsky, D., Harrison, C., Dietz, P., & Wilson, A. (2012, May). Phone as a pixel: enabling ad-hoc, large-scale displays using mobile devices. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 2235-2238). ACM.

[60] Setti, F., Hung, H., & Cristani, M. (2013, July). Group detection in still images by F-formation modeling: A comparative study. In *2013 14th International Workshop on Image Analysis for Multimedia Interactive Services (WIAMIS)*(pp. 1-4). IEEE.

[61] Setti, F., Russell, C., Bassetti, C., & Cristani, M. (2015). F-formation detection: Individuating free-standing conversational groups in images. *PloS one*, *10*(5), e0123783.

[62] Tran, K. N., Bedagkar-Gala, A., Kakadiaris, I. A., & Shah, S. K. (2013, February). Social Cues in Group Formation and Local Interactions for Collective Activity Analysis. In *VISAPP (1)* (pp. 539-548).

[63] Tran, K. N., Gala, A., Kakadiaris, I. A., & Shah, S. K. (2014). Activity analysis in crowded environments using social cues for group discovery and human interaction modeling. *Pattern Recognition Letters*, *44*, 49-57.

[64] Vascon, S., Mequanint, E. Z., Cristani, M., Hung, H., Pelillo, M., & Murino, V. (2014, November). A game-theoretic probabilistic approach for detecting conversational groups. In *Asian Conference on Computer Vision* (pp. 658-675). Springer International Publishing.

[65] Vicon Motion Systems. Website: https://www.vicon.com

[66] Wagner, D., Reitmayr, G., Mulloni, A., Drummond, T., & Schmalstieg, D. (2008, September). Pose tracking from natural features on mobile phones. In *Proceedings of the 7th IEEE/ACM International Symposium on Mixed and Augmented Reality* (pp. 125-134). IEEE Computer Society.

[67] Want, R., Hopper, A., Falcao, V., & Gibbons, J. (1992). The active badge location system. *ACM Transactions on Information Systems (TOIS)*, *10*(1), 91-102.

[68] Wu, M., & Balakrishnan, R. (2003, November). Multi-finger and whole hand gestural interaction techniques for multi-user tabletop displays. In *Proceedings of the 16th annual ACM symposium on User interface software and technology* (pp. 193-202). ACM.

[69] Yamashita, N., Hirata, K., Aoyagi, S., Kuzuoka, H., & Harada, Y. (2008, November). Impact of seating positions on group video communication. In*Proceedings of the 2008 ACM conference on Computer supported cooperative work* (pp. 177-186). ACM.

[70] Yu, T., Lim, S. N., Patwardhan, K., & Krahnstoever, N. (2009, June). Monitoring, recognizing and discovering social networks. In *Computer Vision and Pattern Recognition, 2009. CVPR 2009. IEEE Conference on* (pp. 1462-1469). IEEE.

[71] ZEBRA Corporation. Website: http://www.zebra.com

APPENDIX A

CONSENT FORM

DEPARTMENT OF COMPUTER SCIENCE
UNIVERSITY OF SASKATCHEWAN
INFORMED CONSENT FORM

Research Project: **Object Transfer for Mobile Devices**

Investigators: Dr. Carl Gutwin, Department of Computer Science (966-8646)

Dr. Kevin Stanley, Department of Computer Science (966-6747)

Chengzhao Li, Department of Computer Science

Ashley Coveney, Department of Computer Science

This consent form, a copy of which has been given to you, is only part of the process of informed consent. It should give you the basic idea of what the research is about and what your participation will involve. If you would like more detail about something mentioned here, or information not included here, please ask. Please take the time to read this form carefully and to understand any accompanying information.

This study is concerned with detecting **the optimum way of object transferring for mobile devices**.

The goal of the research is to **test different methods for object transfer for mobile devices**.

The session will require **60** minutes, during which you will be asked to **transfer objects between mobile phones using four different methods of object transfer** in the Human-Computer Interaction Lab at the University of Saskatchewan.

At the end of the session, you will be given more information about the purpose and goals of the study, and there will be time for you to ask questions about the research. As a way of thanking you for your participation and to help compensate you for your time and any travel costs you may have incurred, you will receive a **$10** honorarium at the end of the session.

The data collected from this study will be used in articles for publication in journals and conference proceedings.

As one way of thanking you for your time, we will be pleased to make available to you a summary of the results of this study once they have been compiled (usually within two months).

This summary will outline the research and discuss our findings and recommendations. This summary will be available on the HCI lab's website: **http://www.hci.usask.ca/**

All personal and identifying data will be kept confidential. Confidentiality will be preserved by using pseudonyms in any presentation of textual data in journals or at conferences. The informed consent form and all research data will be kept in a secure location under confidentiality in accordance with University policy for 5 years post publication. Do you have any questions about this aspect of the study?

You are free to withdraw from the study at any time without penalty and without losing any advertised benefits. Withdrawal from the study will not affect your academic status or your access to services at the university. If you withdraw, your data will be deleted from the study and destroyed. Your right to withdraw data from the study will apply until results have been disseminated, data has been pooled, etc. After this, it is possible that some form of research dissemination will have already occurred and it may not be possible to withdraw your data.

Your continued participation should be as informed as your initial consent, so you should feel free to ask for clarification or new information throughout your participation. If you have further questions concerning matters related to this research, please contact:

- Dr. Carl Gutwin, Professor, Dept. of Computer Science, (306) 966-8646, gutwin@cs.usask.ca

Your signature on this form indicates that you have understood to your satisfaction the information regarding participation in the research project and agree to participate as a participant. In no way does this waive your legal rights nor release the investigators, sponsors, or involved institutions from their legal and professional responsibilities. If you have further questions about this study or your rights as a participant, please contact:

- Dr. Carl Gutwin, Professor, Dept. of Computer Science, (306) 966-8646, gutwin@cs.usask.ca
- Research Ethics Office, University of Saskatchewan, (306) 966-2975 or toll free at 888-966-2975.

Participant's signature: ____

Date: ____________________

Investigator's signature: ____

Date: ____________________

A copy of this consent form has been given to you to keep for your records and reference. This research has the ethical approval of the Research Ethics Office at the University of Saskatchewan.

APPENDIX B

QUESTIONNARIES

B.1 Demographic Questionnaire

1. Participant ID:

2. Sex: Male   Female

3. How old are you?

4. Are you a student? Yes   No

5. If yes, what is your field of study?

6. What is your first language?

7. Are you right-handed or left-handed?   Right   Left

8. Of the other three people you are grouped to participate with, how many do you personally know?

9. If you know someone, please list their participant ID

10. On average, how much time do you spend on computers a day?

    Less than 30 minutes

    30 - 60 minutes

    1-2 hours

    2-4 hours

    4-8 hours

    More than 8 hours

11. On average, how much time do you spend on mobile devices a day? (Mobile Devices = mobile phone, tablet)

    Less than 30 minutes

    30 - 60 minutes

    1-2 hours

2-4 hours

4-8 hours

More than 8 hours

12. On average, how much time do you spend on mobile phones a day?

Less than 30 minutes

30 - 60 minutes

1-2 hours

2-4 hours

4-8 hours

More than 8 hours

13. How often do you use touchscreen mobile devices? (Mobile Devices = mobile phone, tablet)

None

Less than 3 hours a week

3-7 hours a week

1-2 hours a day

More than 2 hours a day

14. Please list the type of mobile devices you typically use *

(e.g., Samsung Galaxy S5, iPhone 5, etc.)

15. How do you typically transfer objects (e.g., files, pictures, etc.) from your mobile device to another device? (Check all that apply)

Text Message

Email

Bluetooth

Dropbox/Cloud Service

Near Field Communication

Other:

16. How much time do you spend using word processor, email, or instant messaging?

None

Less than 3 hours a week

3-7 hours a week

1-2 hours a day

More than 2 hours a day

B.2 Post-Condition Questionnaire

1. Participant ID:

2. What was the condition you just encountered?

Cube

Arrow

Wormhole

Tap

3. What was the order that you experienced this condition?

First

Second

Third

Fourth

4. How mentally demanding was the task?

1 2 3 4 5 6 7

Very Low Very High

5. How physically demanding was the task?

1 2 3 4 5 6 7

Very Low Very High

6. How hurried or rushed was the pace of the task?

1 2 3 4 5 6 7

Very Low Very High

7. How successful were you in accomplishing what you were asked to do?

1 2 3 4 5 6 7

Very Low Very High

8. How hard did you have to work to accomplish your level of performance?

1 2 3 4 5 6 7

Very Low Very High

9. How insecure, discouraged, irritated, stressed, and annoyed were you?

1 2 3 4 5 6 7

Very Low Very High

10. Please write any feedback you have for this task or technique:

B.3 Post-Study Questionnaire

1. Participant ID:

2. Rank the conditions, from 1 to 4, in terms of ease of use

   1 = Most Easy

   4 = Least Easy

   * Assign a different number to each condition

   Cube:

   Arrow:

   Wormhole:

   Tap:

3. Rank the conditions, from 1 to 4, in terms of speed

   1 = Most Fast

   4 = Least Fast

   * Assign a different number to each condition

   Cube:

   Arrow:

   Wormhole:

   Tap:

4. Rank the conditions, from 1 to 4, in terms of preference

   1 = Most preferred

   4 = Least preferred

   * Assign a different number to each condition

   Cube:

   Arrow:

   Wormhole:

Tap:

5. Why did you prefer the condition you ranked as 1 over the other conditions?